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 THE ECOLOGICAL RISK ASSESSMENT GE/HOUSATONIC RIVER SITE REST OF RIVER

DCN: GE-060304-ACFL

June 2004

Environmental Remediation Contract
GE/Housatonic River Project
Pittsfield, Massachusetts

Contract No. DACW33-00-D-0006

Task Order 0003

04P-0561-1

RESPONSIVENESS SUMMARY TO THE PEER REVIEW OF THE ECOLOGICAL RISK ASSESSMENT GE/HOUSATONIC RIVER SITE REST OF RIVER

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

Contract No. DACW33-00-D-0006
Task Order No. 0003
DCN: GE-060304-ACFL

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.
Pittsfield, Massachusetts, and West Chester, Pennsylvania

June 2004

W.O. No. 20123.001.096.0735

AUTHORS/CONTRIBUTORS

Tod Delong
Avatar Environmental, LLC
West Chester, Pennsylvania

Scott Ferson, Troy Tucker
Applied Biomathematics, Inc.
Setauket, New York

Dwayne Moore, Roger Breton, Scott Teed, Ryan Thompson
CANTOX Environmental, Inc.
Ottawa, Ontario

Gary Lawrence,
EVS Environment Consultants
Vancouver, British Columbia

Richard McGrath
Sleeman Hanley & DiNitto
Boston, Massachusetts

Francesca Langford
District of Squamish
Squamish, British Columbia

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

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

John Lortie, Bob Roy, Michael Thompson
Woodlot Alternatives, Inc.
Topsham, Maine

PEER REVIEWERS

Valery Forbes
Roskilde University

Thomas W. La Point
University of North Texas

James T. Oris
Miami University of Ohio

Mary Ann Ottinger
University of Maryland

Brad Sample
CH2M HILL

Ralph G. Stahl
DuPont Company

Timothy Thompson
The RETEC Group, Inc.

TABLE OF CONTENTS

Section **Page**

TABLE OF CONTENTS
(Continued)

Section **Page**

LIST OF ACRONYMS

AFDW	ash-free-dry-weight
AhR	aryl hydrocarbon receptor
ANOVA	analysis of variance
BPJ	best professional judgment
CBS	Carolina Biological Supply
COC	contaminant of concern
COPC	contaminant of potential concern
CTDEP	Connecticut Department of Environmental Protection
DELTs	deformities, erosions, lesions and tumors
DL	detection limit
EDC	endocrine disrupting chemical
EPA	U.S. Environmental Protection Agency
EPT	Ephemeroptera, Plecoptera, and Trichoptera
EqP	equilibrium partitioning
ERA	ecological risk assessment
FEL	Fort Environmental Laboratories, Inc.
GE	General Electric Company
GLP	Good Laboratory Practice
HERA	Human and Ecological Risk Assessment
HQ	hazard quotient
HSD	honestly significant difference
IMPGs	Interim Media Protection Goals
LMB	largemouth bass
LOAEL	lowest observed adverse effect level
MATC	maximum acceptable threshold concentration
MCP	Massachusetts Contingency Plan
MDEP	Massachusetts Department of Environmental Protection
MDS	multidimensional scaling
MFO	Mixed Function Oxygenase
MNHESP	Massachusetts Natural Heritage and Endangered Species Program
MSD	Minimum Significant Difference
MSU	Michigan State University
NAPL	non-aqueous phase liquid
NOAA	National Oceanic and Atmospheric Administration
NOAEL	no observed adverse effect level

LIST OF ACRONYMS
(CONTINUED)

PAH	polycyclic aromatic hydrocarbon
PCA	principal components analysis
PRG	Preliminary Remediation Goal
PSA	Primary Study Area
PSDDA	Puget Sound Dredged Disposal Analysis
SETAC	Society of Environmental Toxicology and Chemistry
SIWP	Supplemental Investigation Work Plan
SOP	Standard Operating Procedure
SQT	Sediment Quality Triad
SQV	sediment quality value
SSD	species sensitivity distribution
TDI	total daily intake
TIE	toxicity identification evaluation
tPCB	total PCB
UCL	upper confidence limit
USFWS	U.S. Fish and Wildlife Service
USGS	U.S. Geological Survey
WOE	weight of evidence
YOY	young of the year

INTRODUCTION

This document presents the response from the U.S. Environmental Protection Agency (EPA) to the comments and questions raised by an independent Peer Review Panel following their review of the Ecological Risk Assessment for the GE/Housatonic River Site Rest of River (ERA) released in July 2003. The review was conducted by seven experts in the field of ecological risk assessment. This document, referred to herein as the ERA Responsiveness Summary, has been prepared as part of EPA's obligations under Paragraph 22.c and Appendix J of the comprehensive agreement relating to the cleanup of the General Electric Company (GE) Pittsfield, Massachusetts facility, certain off-site properties, and the Housatonic River (referred to as the "Consent Decree"). The Consent Decree was entered on October 27, 2000, by the United States District Court of Massachusetts - Western Division, located in Springfield, MA.

Under the terms of the Consent Decree, EPA was required to conduct an ecological risk assessment of 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 conclusions of the ecological risk assessment, along with the conclusions from the human health risk assessment that was also conducted by EPA, will be taken into account by GE when developing an Interim Media Protection Goals (IMPG) Proposal that will be submitted to EPA for review.

Prior to the Peer Review, the public comment period provided the opportunity for the public to submit written comments for consideration by the Panel on the risk assessment, within the context of the Peer Review Charge. On January 13-15, 2004, the ERA Peer Review Panel ("Reviewers") met at a public forum in Lenox, MA, to review and discuss the ERA within the framework of the Charge. During this meeting, the members of the public were provided the opportunity to present oral comments to the Panel, and the Panel was able to engage in a question/answer session with the public presenters. The Reviewers subsequently submitted final written comments to EPA's Managing Contractor for the Peer Review, SRA International, Inc., of Arlington, VA. This document is EPA's formal response to the final written Peer Review comments.

APPROACH AND ORGANIZATION OF THIS DOCUMENT

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

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

The first section, termed "Response to General Issues," deals with issues that were raised by a number of Reviewers and/or had broad implications for the ERA. In this first section, EPA has identified 16 General Issues and has provided a Summary of Issue statement for each to frame the technical basis for the issue and to provide an indication of how often the issue was noted by the Reviewers. Each Summary of Issue statement is followed by EPA's response to the General Issue. In many cases, the breadth of the General Issue required a series of responses, each of which is numbered independently. Many of the responses to specific comments from the Reviewers (see below) refer to the responses to the General Issues.

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

RELATIONSHIP OF THE RESPONSIVENESS SUMMARY TO THE ECOLOGICAL RISK ASSESSMENT FOR REST OF RIVER

To better document and integrate the changes to the ERA beyond the scope of the Responsiveness Summary, EPA will issue a revised ERA that will provide additional technical information as described in the responses provided here.

In conclusion, EPA recognizes the hard work and thought that the Reviewers contributed in conducting the Peer Review. As the Reviewers noted, the ERA was a large and complicated document, with many difficult issues being addressed. Although EPA agrees with many of the comments provided by the Reviewers, EPA did not agree with some of the comments; these are documented in the responses and, in such cases, the technical basis for EPA's position is provided. EPA appreciates the effort from the Reviewers in providing their insights and believes that the resulting revised document will benefit greatly from their comments.

RESPONSE TO GENERAL ISSUES

1. DOCUMENT ORGANIZATION, DATA PRESENTATION, AND STATISTICAL METHODOLOGY

SUMMARY OF ISSUE:

In addition to comments directed toward the specific Charge questions, the Reviewers provided a number of suggestions to improve the Ecological Risk Assessment (ERA) document organization and presentation, and also recommended the inclusion of additional types of statistical analyses and presentations.

- Three Reviewers suggested that the ERA would benefit from a reorganization of the document. Currently, the ERA has an executive summary, individual chapters for each Assessment Endpoint, detailed appendices for each endpoint, and several attachments that include the principal investigator reports, including their methods and analyses, and discussions or analyses of information relevant to multiple sections of the ERA. This organization was designed to serve different audiences with varying levels of technical interest in the ERA.

  These Reviewers noted that there was overlap and redundancy between the Assessment Endpoint chapters and their corresponding appendices. The chapters were viewed as providing "too much information for the casual reader and not enough for the interested expert." The Reviewers suggested that the Assessment Endpoint chapters be deleted and the material placed within the corresponding appendices. One Reviewer recommended the inclusion of a detailed road map or data inventory. Other Reviewers stated that the current organization of the ERA is adequate or did not raise this issue in their comments.

- Six Reviewers commented that it was difficult to locate the data that were used in the analyses for each Assessment Endpoint. Suggestions were made to, for example, produce "a series of maps that overlay sampling sites for exposure estimates and sampling sites for the various effects estimates" near the beginning of each Assessment Endpoint chapter and/or appendix. One Reviewer noted that such information would allow a better evaluation of the representativeness and robustness of the data set upon which the exposures are based, particularly in the case of insectivorous birds and omnivorous/carnivorous mammals. It was also suggested that data summary tables be provided for each Assessment Endpoint chapter and/or appendix.

- Five Reviewers noted that more detailed and consistent presentations of the statistical methods used in the ERA are required, rather than simply referring to other documents for statistical details, as was done in places in the current ERA. Five Reviewers commented that a probabilistic presentation of risk, analogous to the wildlife endpoints, would also be useful for aquatic receptors. One Reviewer recommended that other statistical methods, such as species sensitivity distributions or exposure distributions, be considered.

RESPONSE:

1.A. Document Organization

EPA concurs with the suggestions that the Assessment Endpoint chapters in the ERA provide too much detail for the casual reader, and that all of the technical detail supporting the ERA should be included in the appendix chapters. However, EPA believes that casual readers would appreciate summary chapters (i.e., with limited technical detail) in the revised ERA. Therefore, EPA will incorporate the following changes into the revised ERA:

- Provide a "road map" to the document with further discussion that will explain how the summary chapters and appendix chapters are organized and how they relate to each other.

- Shorten the summary chapters to approximately half their current length by deleting much of the technical detail that can be found in the corresponding appendices.

- Review and, as necessary, modify the endpoint appendices to ensure that each is complete and contains the technical details necessary to understand and evaluate the analyses conducted in support of the assessment. It should be noted that in some cases, most notably the GE studies, the information that the Reviewers requested be included may not be available to EPA.

1.B. Presentation of Data

For some endpoints, data were presented in a format similar to that requested by the Reviewers in the ERA. In addition, all data used in the ERA were presented on compact disks provided with the document. However, EPA agrees with the Reviewers' suggestions that additional data presentations would improve the document, and EPA will provide the following in the revised ERA:

- Maps that illustrate the chemical, toxicological, and biological results used in the analyses in the revised ERA, within practical limits of graphical clarity.

- Text and tables that briefly describe each of the data sets (e.g., number and location of samples, sampling design, analytical procedures, analytes, summary statistics, etc.), including screening or transformation steps conducted prior to the analyses.

- For endpoints that rely on data from discrete locations within the Primary Study Area (PSA), a discussion of the representativeness of those data for evaluating exposures throughout the PSA.

1.C. Statistical Methodology

Throughout the ecological risk assessment, EPA used a *p* value of 0.05 in hypothesis testing as the criterion of standard practice for describing an effect as statistically significant. EPA does not intend to change this criterion in the revised ERA. However, it was clear from several of the comments from the Reviewers (e.g., see General Issues 15.B and 16.B), that too much emphasis was placed on the "bright line" that *p* must be less than 0.05 for the observed effects to even be considered in the ERA. EPA believes that it is appropriate to consider effects in ecological risk assessments even when the *p* value is > 0.05 for the following reasons.

The original designation of the p value of 0.05 was to achieve a different objective from what EPA is striving to achieve in an ecological risk assessment.

The underlying philosophical concept for hypothesis testing is that scientific advances come from framing clear hypotheses that are subject to rejection through experimentation (Platt 1964; Popper 1968). In most scientific endeavors, "the burden of proof is on the scientist who posits a new natural phenomenon" (Suter 1996). As a result, most hypothesis tests set the Type I error rate (α) for rejecting a true null hypothesis to be low (i.e., α or $p = 0.05$) while allowing for a much higher Type II error rate (β) for accepting a false null hypothesis (Sokal and Rohlf 1981). This practice was established because the cost of a Type II error was considered to be relatively small (e.g., scientists must perform better-designed experiments or posit a more plausible hypothesis) while the cost of a Type I error was higher (e.g., other scientists waste funds pursuing research on a new natural phenomenon that does not actually exist, credibility of scientific community is undermined). As stated by Suter (1996), "Therefore, because science is a cumulative enterprise that builds on a foundation of reliable knowledge, Type I error is carefully controlled to 5% or lower, but Type II error is either ignored or allowed to assume higher values (Parkhurst 1990). As a result, hypothesis testing has a strong and explicit bias in favor of the null hypothesis." Thus the burden of proof is placed on the investigator and on whether the data represent a preponderance of evidence sufficient to reject the null hypothesis. The imposition of this bias on ecological risk assessment has the unfortunate consequence that adverse effects to the environment are difficult to demonstrate even when they are or may be occurring (Mapstone 1995; Suter 1996; and Heath 1990). In such circumstances, a risk may not be detected even though in fact there are adverse effects to environmental receptors, because ecological responses to environmental contaminants in the field are difficult to detect because they are often nonlinear, non-normal, and heteroscedastic, and the sampling effort that could be devoted to the question is limited (Legendre and Legendre 1998).

Lack of a finding of statistical significance does not imply that no effects occurred

Moore and Caux (1997) analyzed 24 data sets that had an adequate model fit to the concentration-response relationship and found that 76.9% of the NOAELs and 100% of the LOAELs were higher than the corresponding point estimates for 10% effect from the best-fit model equation (Moore and Caux 1997). Similar results have been obtained by EPA (1991) and Hoeckstra and Van Ewijk (1993). Oris and Bailer (1993) found that minimum detectable mean differences between controls and treatments ranged from 31 to 100% inhibition for Ceriodaphnia reproduction in tests following the standard EPA protocol. Together these and other results indicate that effects below a magnitude of 10% are rarely detected as significant, and that more often than not, effects must be greater than 30% to be detected as significant.

The above results were for tightly-controlled laboratory toxicity tests that followed standardized toxicity testing protocols. Field studies generally have much lower statistical power than laboratory toxicity tests because of low sample size, high background variation, and other factors. Thus, effects much greater than 30% may be required before such effects become statistically significant at p=0.05. EPA (1997) gives the option of letting the risk assessor adjust the p value upward (e.g., p=0.1, although this should be done a priori) to ensure that effects of biological significance have a higher potential of being detected as statistically significant.

Relevance of effects is not considered

Hypothesis testing makes a determination regarding statistical significance on the basis of a process that combines magnitude of the effect with the variance among replicates (Suter 1996). Thus, a statistically significant effect may be the result of an effect of high magnitude, low intra-treatment variance, or both. The biological or ecological significance of the effect are not considered. Consider the case where a 40% or greater effect is required for an effect to be statistically significant at p of 0.05. Strict reliance on this p value in order for the effect to be considered in the ERA would de facto mean an assumption is made that effect levels lower than 40% are not of significance regardless of biological, ecological or societal concerns. A more appropriate approach would be to consider the magnitude and biological relevance of observed effects as part of the overall weight-of-evidence assessment, even if those effects would only be statistically significant if p was greater than 0.05 (e.g., p=0.1). The fact that the effects were not statistically significant, and an explanation of the possible reasons for this result, such as high background variation or other factors, should be provided as part of the overall discussion on sources of uncertainty.

Summary

In the revised ERA, EPA will:

- Continue to use p=0.05 as the criterion for determining statistical significance;
- Include the results of studies for effects that may be biologically significant in the overall risk characterization, even if these results are not statistically significant with a p of 0.05; and
- Discuss in the sources of uncertainty section any effects considered in the weight-of-evidence assessment that were not shown to be statistically significant at a p of 0.05.

Although the statistical methods were included for the most part in the investigator reports, EPA concurs with the suggestion by the Reviewers to provide detailed descriptions of the statistical methods used, including their assumptions and limitations, in the revised ERA endpoint appendices. Where appropriate (e.g., for hypothesis tests), statistical significance, power and effect size will be reported.

EPA agrees that the consideration of additional statistical techniques may enhance the analyses presented in the ERA and improve the risk characterization, and will consider the use of alternative presentation formats in the revised ERA, such as:

- Cumulative probability plots of exposure, against which one or more effects thresholds can be compared.
- Regression analyses.
- Species sensitivity distribution (SSD) plots.

- Other tabular or graphical presentations that link probability of exposure and probability of effect.

2. ASSESSMENT ENDPOINT DEVELOPMENT

SUMMARY OF ISSUE:

Five Reviewers commented on the definition of the Assessment Endpoints evaluated in the ERA, generally indicating that the endpoints should be redefined to better reflect the specific level of organization that is of interest for each endpoint. Two Reviewers noted that the level of ecological organization of the various Assessment Endpoints is variable and results in difficulties in assigning priorities, as well as creating redundancy. One Reviewer recommended that the societal value of the receptors should be considered in defining the Assessment Endpoints, and another recommended that the Assessment Endpoints should be reworded to reflect more specifically those attributes that are most worthy of protection. Finally, one Reviewer noted that some endpoints were not stated in the form of specific, testable hypotheses.

RESPONSE:

EPA recognizes the comments expressed by the Reviewers regarding the wording of the Assessment Endpoints. However, EPA does not believe that it is appropriate to retroactively redefine the Assessment Endpoints after completing the risk assessment. The development and wording of the Assessment Endpoints for the Housatonic River ERA have a long history, including the involvement of the Massachusetts Department of Environmental Protection (MDEP), Connecticut Department of Environmental Protection (CTDEP), Mass Wildlife, National Oceanic and Atmospheric Administration (NOAA), U.S. Fish and Wildlife Service (FWS), EPA, and General Electric Company (GE), some of which preceded the Agreement in Principle between the parties established in 1998. EPA will provide additional discussion in the Problem Formulation and Risk Characterization sections to further clarify the intent of the endpoints, where possible.

3. SCREENING OF COPCs and DETERMINATION OF COCs

SUMMARY OF ISSUE:

Most Reviewers generally agreed that the contaminants of potential concern (COPC) selection process was conducted appropriately, some noting that the COPC selection was “extensive,” “thorough,” or, in one case, that the process would have passed for a baseline risk assessment at other sites. Four Reviewers, however, made some specific comments, including that the selection was not conservative, could be better explained, or needed additional discussion to address the exclusion of several contaminants detected in sediment and biological tissues.

With reference to the aquatic endpoints, three Reviewers noted that contaminants of concern (COCs) other than PCBs (e.g., polycyclic aromatic hydrocarbons [PAHs] and/or mercury) could have been responsible for some of the observed effects.

RESPONSE:

EPA agrees that more discussion would be helpful regarding the elimination of certain pesticides, PAHs, and mercury as COPCs for several Assessment Endpoints. An expanded discussion of the COPC selection process as it pertains to the elimination of COPCs will be provided in Appendix B (Pre-ERA).

Where appropriate, additional justification for the further reduction of the COPC list for specific Assessment Endpoints will be included in the associated endpoint section. Some PAHs and metals (including mercury) were retained as COPCs following the Pre-ERA screening process (see Tables B-145 through B-148). COPCs remaining at the conclusion of the three-tier analysis performed in the Pre-ERA were further screened in the individual Assessment Endpoint evaluations on a medium-specific basis, using more detailed and receptor-specific criteria, resulting in the definitions of Assessment Endpoint species COCs.

In the benthic invertebrate (Section 3.2.1) and amphibian (Section 4.3.1) assessments: PCBs, PAHs, and several metals (including mercury) were retained as COCs; and PCBs and PAHs were retained as COCs for the assessment of fish (Section 5.2.1). There was no relationship between concentrations of PAHs and metals in sediments and toxicity and benthic community structure.

For amphibians, both PAHs and metals (including mercury) were either close to background concentrations and/or only slightly exceeded screening effects concentrations. PCBs, however, exceeded sediment quality benchmarks by 1 to 2 orders of magnitude at most amphibian sampling sites evaluated in the PSA.

Sediment PAH concentrations were considered in the fish risk assessment and threshold values for three specific PAHs and for total PAH were developed. For most areas of the PSA, the risks to fish due to sediment concentrations of PAHs were low. For example, the HQs for total PAH, based on median sediment concentrations and using the 10 mg/kg literature-based effects threshold, were below 1.0 in all reaches (5A, 5B, 5C, and 6). HQs based on mean concentrations were slightly higher; exceeding 1.0 in Reach 5A and 5C, but were below 3 in all reaches in the PSA. Furthermore, because the assessment for PAHs in sediment relied on an effects threshold derived for bottom fish (i.e., brown bullhead), the derived HQs are expected to overestimate risks to fish with a water column-based life history. The latter would have lower exposures due to both PAH metabolism in teleost fish (and consequent limitation on trophic transfer of PAHs to predator fish) and reduced contact with the sediment bed.

Because the HQ was derived using literature-based benchmarks and an effects threshold for the most highly exposed species, the HQ assessment is considered to be conservative with associated uncertainty, and that small exceedances of HQ = 1.0 in this screening step should not be interpreted as evidence of harm.

Metals, including mercury, were eliminated from consideration in the COC screening for fish because the sediment concentrations measured in the PSA were below applicable benchmarks and/or concentrations measured in reference areas (see ERA Attachment F.1).

Based on the analyses described above, PCBs were implicated as the dominant causative agent for risk for all aquatic endpoints and much of the risk characterization discussion for these

Assessment Endpoints focused on PCB-related impacts. EPA will review the non-PCB COC selection and evaluation process presented in each Assessment Endpoint section and will provide additional clarification where appropriate.

4. USE OF INFORMATION FROM FIELD STUDIES

SUMMARY OF ISSUE:

Four Reviewers commented that several studies described in the Supplemental Investigation Work Plan (SIWP) were not described or mentioned in the ERA and that the data from the studies did not appear to be included in the ERA. One Reviewer recommended that an explanation be provided of why dragonflies and waterfowl, specifically, were not included as receptors or representative species. The studies specifically noted included:

- Dragonfly Study (SIWP App. A.7)
- Mussel Survey and Caged Mussel Study (SIWP Apps. A.8 and A. 16, respectively)
- Bluegill Toxicity Study (SIWP App. A.22, in part)
- Sampling and Analysis of Crayfish (SIWP App. A.17)
- Waterfowl Collection and Tissue Sampling (SIWP App. A.24)

It was recommended that more explicit recognition of these studies be made in the revised ERA.

RESPONSE:

Information from all of the studies listed in the SIWP was considered in the ERA to describe the food chain and develop the conceptual model, to identify potential species at risk, or to characterize risk to a specific species. However, not all studies described in the SIWP were intended to be components of an Assessment Endpoint. Additional text will be added to the ERA explaining how all of the studies cited in the SIWP were used.

In the risk assessment, certain species were identified as being representative of the varying habitats and exposure pathways present in the PSA. The process outlined by Suter et al. (2000) was used to choose representative species for inclusion in the ERA. For some of these representative species, site-specific studies were conducted to evaluate risk. As discussed in the ERA, risk to other species observed or expected to be present in the Rest of River, which were ecologically similar to those included in the risk assessment, were described in the Ecological Significance sections of each appendix.

EPA agrees that better explanation of how these supporting studies were considered in the ERA would improve transparency. The use and/or disposition of the specific studies in question are summarized below, and will be documented in the revised ERA.

4.A. Dragonfly Study

Dragonfly surveys were conducted to determine if any threatened or endangered species were present in the river, based on the recommendation from a reviewer of the SIWP. Site-specific toxicity testing was not performed on dragonflies in the PSA because a number of other studies on representative invertebrate species were planned. Risks to dragonflies and other invertebrates in the river were characterized based on risks to representative invertebrate species that were studied more intensively and considered representative of the local benthic community in its entirety.

4.B. Mussel Survey and Caged Mussel Study

A survey of freshwater mussels and a study of the survival and physiology of freshwater mussels were included in the SIWP. The mussel survey was conducted for the same purpose as the dragonfly survey discussed above in the response to General Issue 4.A, and the results were used similarly. The results of the caged mussel study were not incorporated into the ERA because the study was terminated prior to completion due to changes in river conditions that interfered with the deployment and monitoring of animals (bedload during a major storm event buried a number of cages). Accordingly, no usable biological data were obtained, and the study was limited to chemistry measurements made in the vicinity of the mussel cages. The latter data were used in the ERA.

4.C. Bluegill Toxicity Study

Evaluation of the reproduction of bluegill was included in the Work Plan for the Phase I Fish Toxicity Study. The experiment was initiated as planned and indigenous bluegill were shipped to the U.S. Geological Survey (USGS) Columbia laboratory. The deployment of the bluegill was delayed until completion of the largemouth bass spawning; subsequently, the location and collection of bluegill eggs in the experimental ponds at the USGS facility was largely unsuccessful. The Principal Investigators identified the most probable reason for the lack of spawning success to be the extended holding times in tanks prior to deployment in the ponds (2 months) which may have interfered with gonadal maturation in the fish.

4.D. Sampling and Analysis of Crayfish

The objective of the crayfish study was to document contaminant concentrations in crayfish tissue, both upstream of and within the PSA. The data were collected primarily to support the food-chain modeling study, and were also used as a dietary component for the ERA receptors, when appropriate.

4.E. Waterfowl Collection and Tissue Sampling

Waterfowl data were collected from trapping locations in Threemile Pond (n=20), Reaches 5C and 5D (n=18) and Reach 6 (n=7) in 1998. The collection in Reaches 5C, 5D, and 6 also included some birds obtained during the Mass Wildlife banding efforts conducted in the same locations shortly before the trapping. Two species were collected, mallards (*Anas platyrhynchos*) and wood ducks (*Aix sponsa*). Total PCB (tPCB) and TEQ concentrations were determined in all samples. The resulting data were used in the exposure assessments for bald eagles and mink because birds (primarily waterfowl) are significant components of the diet of both species.

In response to the Reviewer comments, EPA will add a waterfowl receptor to the revised ERA. The waterfowl tissue data will be used to estimate exposure in adults. Contaminant concentrations in wood duck eggs (via maternal transfer) were measured in a limited sampling effort conducted in May 2004. These data will be incorporated into the revised ERA.

5. EXTENT OF SAMPLING AND CHARACTERIZATION IN CONNECTICUT

SUMMARY OF ISSUE:

Four Reviewers commented on the reduced intensity of sampling in Connecticut (CT). One of these Reviewers recommended that the reduced sampling in CT needed to be better justified or, if not, that the contamination in CT should be better characterized. Three of these Reviewers specifically noted the potential for accumulation of more highly contaminated sediment behind the large dams in CT and recommended that EPA consider sampling behind the dams; or, if data were already available from behind the dams, that the data be presented in the ERA.

RESPONSE:

EPA does not agree that the characterization of the river in CT was superficial or inadequate. The Ecological Characterization conducted for the EPA Rest of River study and included in the ERA as Appendix A includes a separate volume (Appendix A.2) that provides extensive physical, ecological, and cultural information for the river downstream of Woods Pond, including the full extent of the river into and through the seven reaches (Reaches 10 to 16) in CT (Reach 17 is tidal and has other sources of PCB contamination, so is excluded from the Rest of River study).

A robust set of historical and recent data documenting contaminant concentrations in water, sediment, and biota in CT is available and was used in the ERA. The extent of sampling of all media in CT, which included sampling conducted both prior to and during the EPA Rest of River study, was commensurate with the extent of contamination, and is adequate to estimate the risks to ecological receptors in CT. The data from CT include recent sampling conducted by EPA, which specifically included confirmation of contaminant concentrations in sediment behind dams.

An overview and summary evaluation of the historical and recent sediment and biota (tissue) data are presented below. These data were presented and discussed in detail in the RFI Report (BBL and QEA 2003) and sediment and surface water data were also presented in the ERA (Section 2.5.2). The discussion in the revised ERA will be expanded to include a presentation of the biota (fish and benthic invertebrate) data from CT.

Extent of Sampling in CT

Sediment samples were collected historically (since approximately 1980) from the CT reaches of the river, along with a number of water samples and biota tissue samples. Sampling of fish (trout and smallmouth bass) and invertebrates in the West Cornwall area (Reach 12) has been conducted biannually since 1984 by GE.

Sediment sampling to support the Rest of River study was conducted in an iterative fashion. Historical data were reviewed in designing the initial phases of this program; subsequent phases were then implemented after a review of these data. Sampling in Reach 9 conducted as part of the Rest of River study documented substantially decreased sediment PCB concentrations from Rising Pond dam downstream to the MA/CT state line, a distance of 24 miles. Samples of floodplain soil collected from Reach 9 confirmed that this marked decline was also reflected in tPCB concentrations in the floodplain.

Consistent with the iterative sampling strategy discussed above, these data were reviewed and additional extensive sediment sampling in CT was judged not necessary due to the very low concentrations both historically in CT and currently in the 24-mile reach (Reach 9) immediately upstream in MA. Additional floodplain sampling was similarly judged not necessary due to the very limited floodplain contamination in Reach 9; in addition, there is very limited floodplain in CT due to the change in topography. However, evaluation of the historical sediment data from CT did indicate a number of areas where additional data would be useful to address some specific objectives.

In response, EPA conducted a supplemental sediment sampling program in 2001, in which additional samples were collected and analyzed for tPCBs, grain size, and TOC, thus providing a combined data base of over 500 sediment samples from CT (excluding nearly 100 additional samples in Reach 17). The supplemental sampling program was designed to provide data to:

- Determine PCB concentrations in areas not previously sampled.
- Verify the spatial/temporal distribution of PCBs, including the distribution of PCBs upstream of dams.
- Determine PCB concentrations in areas where public access to the river is available.

PCB Concentrations in CT Sediment

Of the samples analyzed from the EPA 2001 sampling program, only 15% had tPCB concentrations above the laboratory detection limit (approximately 0.02 mg/kg). The highest concentration of tPCB detected was 1.2 mg/kg, in sediment (0.5 to 1 ft interval) from Reach 14 (Lake Lillinonah). All other samples were less than 1 mg/kg tPCB.

Sediment tPCB concentrations from all reaches of the Housatonic River are presented in Figure 1. As shown in this figure, tPCB concentrations decrease markedly downstream of Woods Pond and decrease farther downstream of Rising Pond through Reach 9, the last reach in MA. Downstream of the MA/CT state line and into Connecticut, PCB concentrations remain low until Reach 17, which is tidal and receives other sources of PCBs.

Figures 2 through 4 present the tPCB concentrations for all samples (surficial and at depth), including data from both the historical and 2001 EPA sampling, by river mile for Reaches 12, 14, and 15, which are the only CT reaches with any sediment PCB concentration detected, either historically or in 2001, in excess of 2.0 mg/kg.

All of these higher concentrations (>2.0 mg/kg) are from the historical sampling program. The highest PCB concentrations detected in sediment from CT include 5.9 mg/kg (Reach 15 at 2 feet below river bed surface) and 8.2 mg/kg (Reach 14 at a depth of between 1.0-1.5 feet).

Figures 5 through 7 present tPCB concentrations at depth in sediment, including data from both the historical and 2001 EPA sampling, for Reaches 12, 14, and 15. As these figures show, tPCB concentrations in the more recent sampling were considerably lower than previous years, and the highest concentrations were approximately 2 feet deep in the sediment.

PCB Concentrations in CT Biota

Several hundred fillet samples from several species of fish have been collected from CT and analyzed for tPCBs over approximately the past 25 years. These data demonstrate a clear decrease in tPCB tissue concentrations in CT fish over this period, with most of the decrease occurring prior to the early 1990s. Figure 8 presents data for tPCBs in smallmouth bass fillet samples collected from 1979 to 1996 from three river reaches in CT; these results are typical of those for all species and all reaches. The decline in tissue concentrations of tPCBs from the initiation of sampling through the early 1990s is evident; however, the lower concentrations remain unchanged since that time.

SUMMARY

The extensive data record for CT indicates that the residual tPCB contamination in sediment is at least one to two orders of magnitude lower than in the PSA and in other isolated areas in MA. There is a clear indication that concentrations in all media in CT have decreased from the initiation of the data record until the early 1990s. These lower concentrations and the decreases over time until the early 1990s, are reflected in the fish tissue concentrations. EPA believes these results are sufficient to support the estimates developed in the ERA regarding risks in CT.

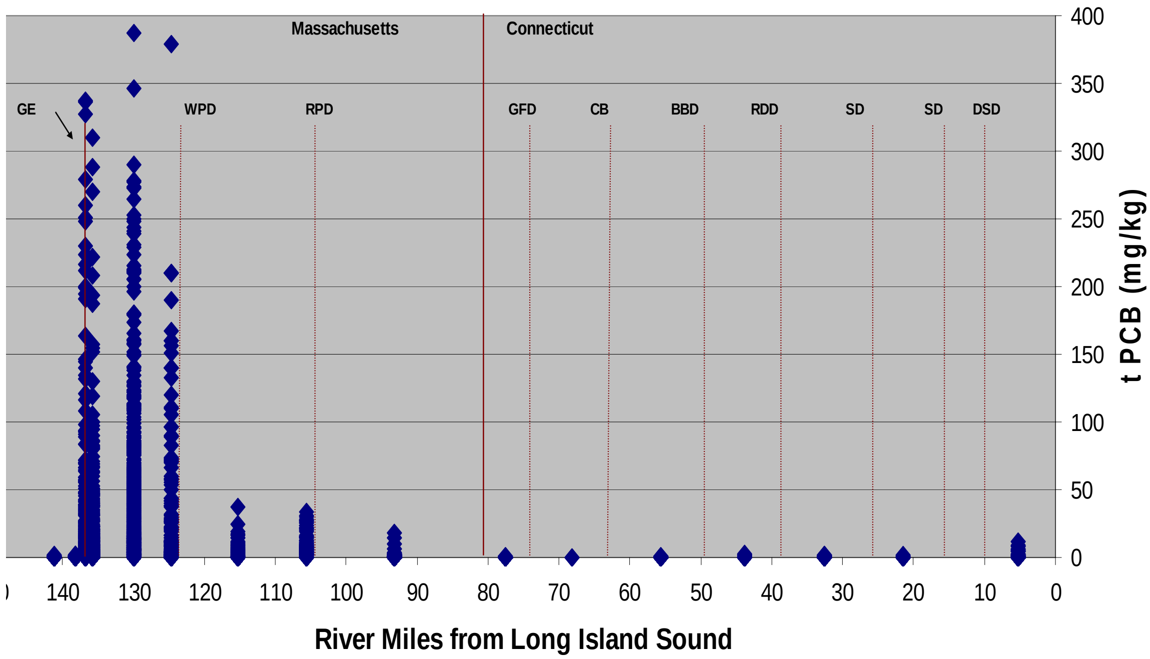


Figure 1 Sediment tPCB Concentrations (1990 to 2002 – top 3 feet)

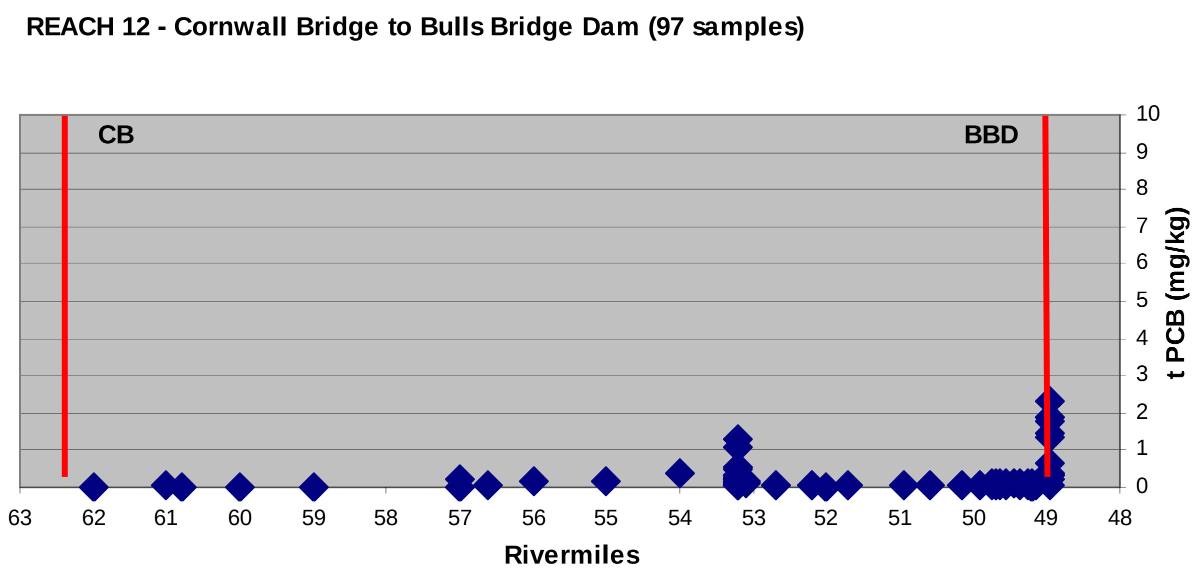


Figure 2 Sediment PCB Concentrations – Reach 12

REACH 14 - Bleachery Dam to Shepaug Dam (153 samples)

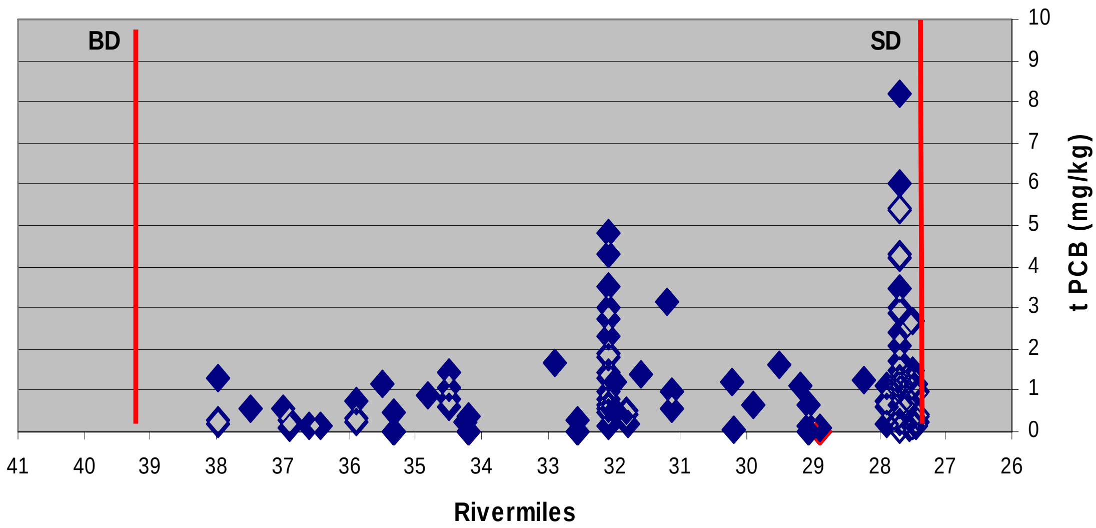


Figure 3 Sediment PCB Concentrations – Reach 14

REACH 15 - Shepaug Dam to Stevenson Dam (125 samples)

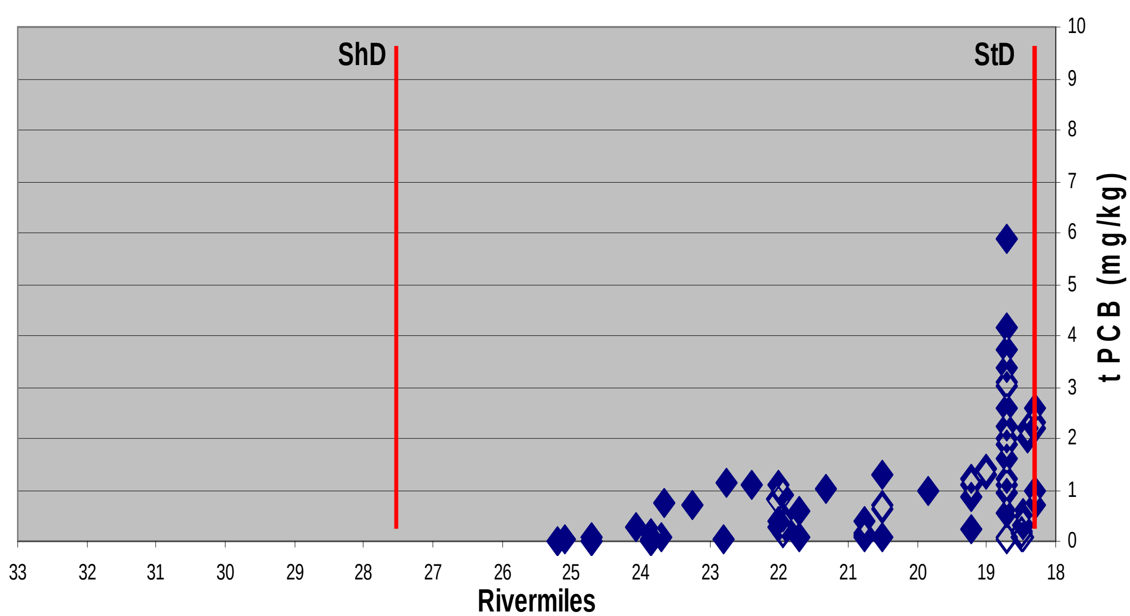


Figure 4 Sediment PCB Concentrations – Reach 15

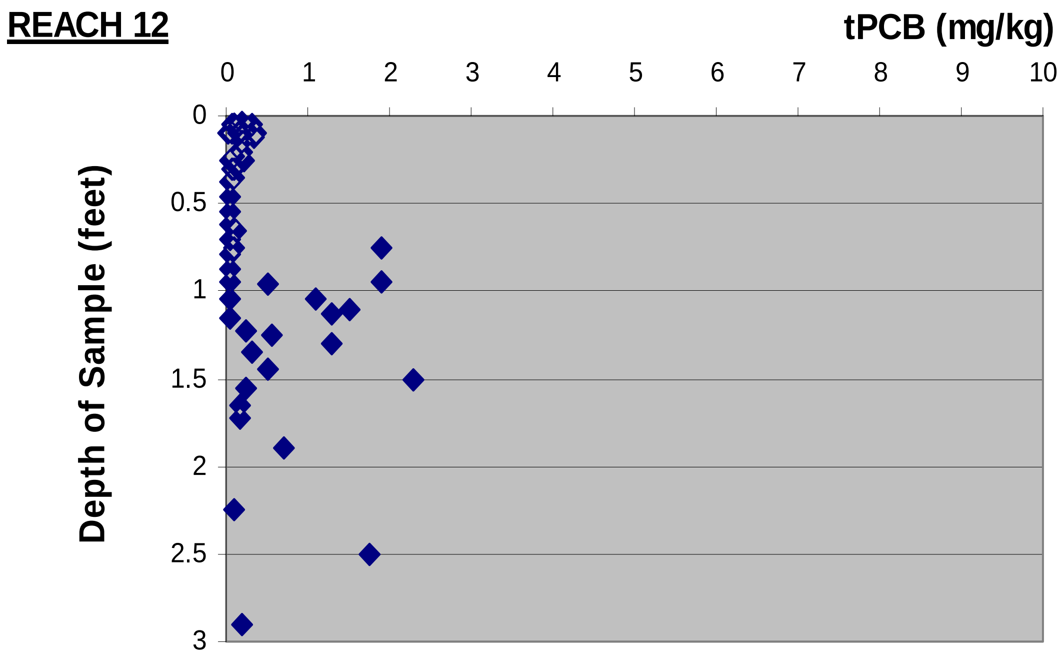


Figure 5 Sediment PCB Concentrations by Depth – Reach 12 (all data)

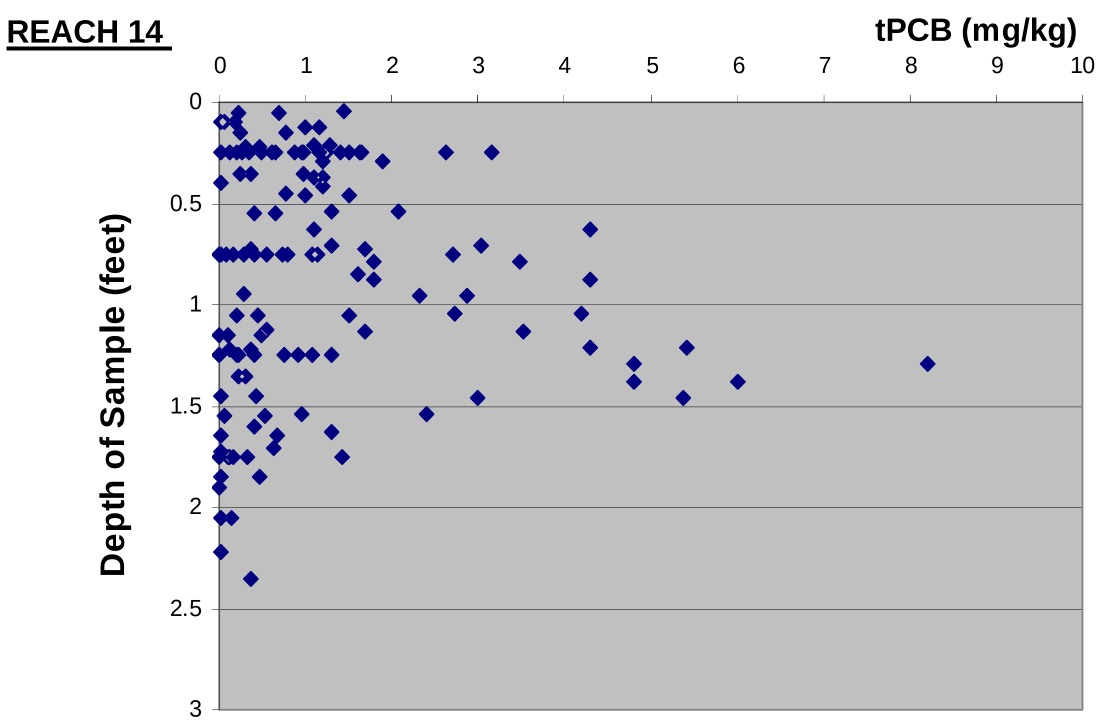


Figure 6 Sediment PCB Concentrations by Depth – Lake Lillinonah, Reach 14 (all data)

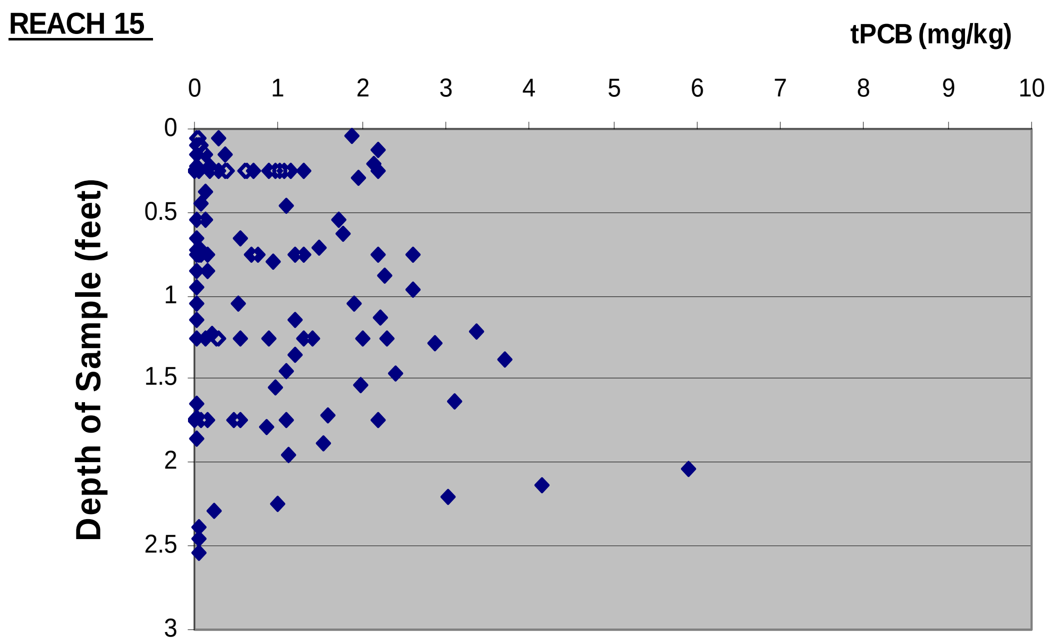


Figure 7 Sediment PCB Concentrations by Depth – Lake Zoar, Reach 15 (all data)

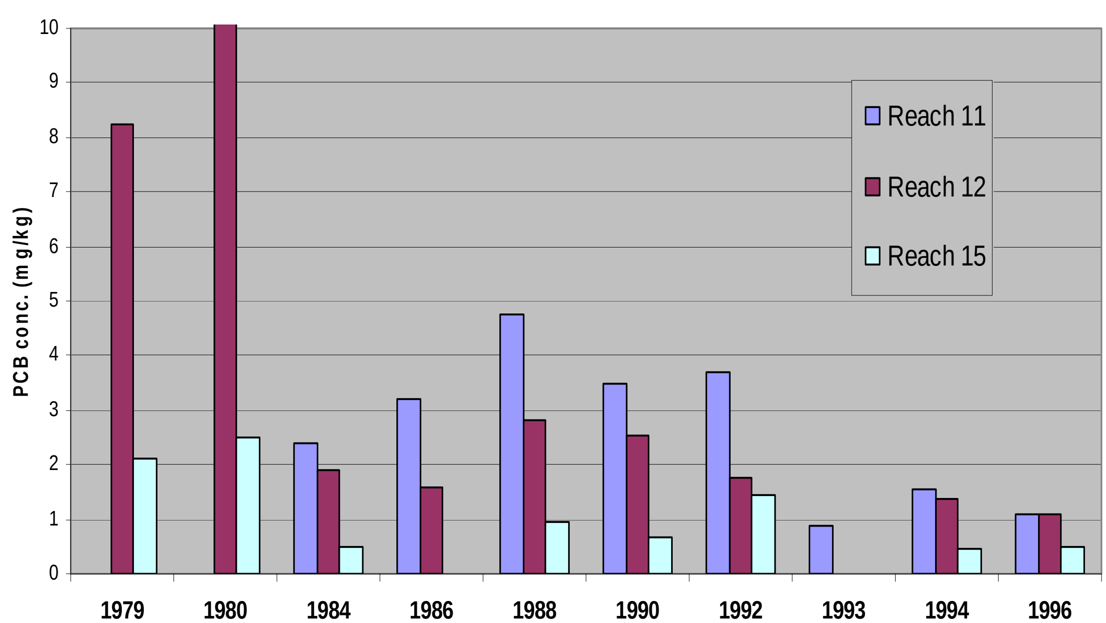


Figure 8 Connecticut Smallmouth Bass (Fillet) Mean PCB Concentrations 1979 - 1996

6. DEVELOPMENT OF MATCs

SUMMARY OF ISSUE:

Five Reviewers recommended the establishment of decision rules for derivation of the maximum acceptable threshold concentrations (MATCs). Although the proposed decision rules were developed for the benthic invertebrate MATC, one Reviewer commented that the rules should apply to the calculation of all MATCs for all Assessment Endpoints. The specific recommendations of the Reviewers are summarized below:

- Separate acute and chronic MATCs – Four Reviewers recommended that, data permitting, acute and chronic endpoints be considered separately in MATC derivation. Of the four, two Reviewers further indicated that the final MATC should be weighted toward results of chronic endpoints.

- Eliminate redundant endpoints – Three Reviewers commented that MATCs in the ERA were calculated based on multiple measurements of the same type of effect for the same test species (e.g., dry-weight vs. ash-free dry-weight). It was recommended that only a single value be used for each effect type for each test species per MATC. One Reviewer recommended that only one endpoint be used per species, either the lowest observed effect or the geometric mean of the response for the endpoints.

- Use "most synoptic data" – Five Reviewers commented on the exposure data used in the development of MATCs, particularly for the benthic invertebrate endpoint. One Reviewer commented that an "error" had been made, and three Reviewers stated that the MATC derivation should emphasize dose-response identified using the "most synoptic" exposures for the toxicity tests.

- Use only tests endpoints that exhibit dose-response relationships – Three Reviewers suggested that the MATC derivation should consider only test endpoints that demonstrate dose-response relationships. Details on what constituted an acceptable dose-response relationship were not provided.

- Use only sediment-relevant receptors – Three Reviewers questioned the suitability of cladoceran species (i.e., *Daphnia* and *Ceriodaphnia*) for evaluating sediment toxicity.

- Use all available data – Three Reviewers indicated that MATC values be based on all acceptable toxicity values, not just the six lowest values. This comment was specific to the benthic invertebrate ERA, for which MATCs were originally calculated based on the lowest six toxicity test values.

- Comparison of MATC to reference – Three Reviewers commented that if the derived MATC for benthic invertebrates is equal to or lower than the concentration at reference sites, then the reference site concentration should be used as the MATC.

The Reviewers provided differing opinions regarding the conservatism inherent in the MATCs. For example, for benthic invertebrates, one Reviewer stated that "the MATC of 3 mg/kg tPCB is likely overly conservative," but another Reviewer stated that "given the complexity of this class of receptor, more conservative MATCs are appropriate."

RESPONSE:

EPA agrees that the application of a consistent set of decision rules in the derivation of the MATCs will improve the ERA, and agrees with most, but not all, of the specific recommendations summarized above. Each recommendation is addressed individually in the following subsections.

6.A. Separate Acute and Chronic Endpoints

No Reviewer defined the terms "acute" and "chronic" in their written comments. However, based on a review of the videotaped discussion between the Reviewers at the Peer Review Meeting, EPA interprets the recommendation to mean separation of the endpoints based on the duration of contaminant exposure in relation to organism lifespan, rather than based on whether mortality was the endpoint. Based upon this interpretation, the following definitions will be used (EPA 2004):

- Chronic Toxicity: The capacity of a substance to cause long-term poisonous health effects in humans, animals, fish, and other organisms.

- Acute Toxicity: The ability of a substance to cause severe biological harm or death soon after a single exposure or dose. Also, any poisonous effect resulting from a single short-term exposure to a toxic substance.

Based on the above definitions, EPA believes that distinguishing between acute and chronic endpoints is appropriate. This will be addressed in the revised ERA as follows:

- Benthic Invertebrates – The in situ exposures will be considered acute, because they were 48 h, 7 days, and 10 days in duration. The laboratory exposures (including survival, growth, and reproduction endpoints) will be considered chronic, because they were over 40 days in duration and covered the reproductive period. Separate MATCs will be developed for these two exposure durations.

- Amphibians – Chronic MATCs were developed using chronic endpoints for wood frogs, based on exposures conducted through the reproductive period. No acute endpoints were tested; therefore, calculation of an acute MATC is not possible.

- Fish – The fish toxicity studies are considered chronic toxicity studies, following the above definitions. No acute endpoints were tested; therefore, calculation of an acute MATC is not possible. Separate MATCs will not be developed for lethal and sublethal endpoints, since the deformities observed in fry are considered sufficient to impair the probability of survival to the adult stage. The purpose of the toxicity studies was to evaluate the effects of COCs on early lifestage development and reproduction; discrimination between mortality and major pathologies is not appropriate because there is not the biological or toxicological foundation for separating these effect types at this lifestage.

- Wildlife – The wildlife MATCs were intended to be protective for a chronic exposure through the diet. In the case of the mustelids (mink and river otter), the long-term feeding study used in the ERA (i.e., Bursian et al. 2002) will be the basis for the MATC in the revised ERA. For the bald eagle, the field studies linking dietary exposure levels to effects from the literature that were used in the current ERA will be the basis for the MATC in the revised ERA. MATCs for other wildlife species were not calculated, because either the risks in the PSA were shown to be low (e.g., tree swallow, American robin), or the risks were more uncertain because species-specific data or data from a reasonable surrogate were not available (e.g., belted kingfisher, fox, osprey).

6.B. Eliminate Redundant Endpoints

EPA generally agrees with the Reviewers' comments and will address them in the revised ERA as follows:

- Benthic Invertebrates – Only one value for effect duration will be used. For example, where 28-d, 35-d, and 42-d results are available for an endpoint, only one of the values will be applied in MATC derivation. Results for all three durations will be reported in the ERA effects assessment as supporting lines of evidence, because these measurements represent different test replicates. Where both ash-free-dry-weight (AFDW) and dry-weight growth measurements exist, the MATC will consider only the AFDW, because this measure has the potential to remove bias caused by material in gut contents that does not reflect somatic growth (Sibley et al. 1997).

- Amphibians – No change required. If gonadal and overall malformation rates had been used to develop different MATCs, there would have been potential for redundancy; however, this was not done in the ERA.

- Fish – For some fish toxicity tests, results of multiple trials are available. These trials represent different batches of eggs, that may have different sensitivities to contaminants introducing another source of variability. Furthermore, multiple treatments are available representing different exposure durations throughout development. In the revised ERA, the MATC derivation will continue to be based on multiple measurements (trials and test durations) for each species. However, the ERA will more clearly describe the number of endpoints included for each species, and discuss any potential for bias in the MATC estimates.

- Wildlife – Redundant endpoints were not used in deriving the MATCs in the current ERA, nor will they be used in the revised ERA.

6.C. Use of Most Synoptic Data

The dose-response evaluation for the benthic invertebrate ERA was conducted using <u>both</u> the "most synoptic data" and the "median of data," and the results from both analyses were presented in the ERA. Both approaches were used because EPA recognized the same issues raised by the Reviewers. Attachment D.5 to the ERA (Alternative Concentration-Response Assessment for PCB and Toxicity Test Endpoints) provided a detailed analysis using only the "most synoptic data." In the development of a final MATC, there was general concordance between the MATCs generated from the two data processing methods. EPA will revise the ERA as follows:

- Benthic Invertebrates – EPA believes that consideration of both data analysis approaches is appropriate, given the high variability in individual measurements of sediment PCB concentrations. However, in recognition of the recommendation of several Reviewers that the "most synoptic" data are the more appropriate exposure measures, the revised ERA will bring the Attachment D.5 material forward to the main Appendix D text, and move the analysis using the median concentration approach to the attachment.

- Amphibians – No changes required. The exposure data were synoptic as they were either sediment PCB concentrations from the mesocosms where the frog rearing occurred (Phase I), or they were sediment PCB concentrations from the pools where wood frogs were reared in the field pools (Phase II and III).

- Fish and Wildlife – No changes required. For these mobile receptors, the exposure data were treated on a spatial scale commensurate with the home ranges of the receptors, or were based on controlled laboratory studies.

6.D. Only Use Endpoints with "Dose-Response" Relationship

EPA agrees that a dose-response relationship should be established to allow derivation of an MATC. However, EPA does not believe that every treatment and endpoint used in MATC derivation must have a smooth and continuous (i.e., monotonically increasing) dose-response shape. There are many explanations for deviations from continuous dose-response curves, including uncertainty in the exposure concentrations, natural variability, the potential for homeostasis, and other factors.

Deviations from perfectly smooth dose-response relationships are common in toxicological assessments. EPA protocols explicitly acknowledge the need to "smooth" discontinuous data. For example, EPA (2002) refers to derivation of toxicity thresholds using the Spearman-Karber method, and states: "if the response proportions are not monotonically non-decreasing with increasing concentration (constant or steadily increasing with concentration), the data are smoothed."

Based on Reviewer comments, EPA will implement the following modifications in the revised ERA:

- Benthic Invertebrates – Dose-response relationships for all toxicity endpoints will be examined graphically and statistically. Criteria for inclusion of an endpoint in MATC derivation for benthic invertebrates will include the following:

  - Negative control performance will not be considered in MATC derivations.

  - A statistically significant difference between field reference sediment treatment and at least one exposure treatment must be observed.

  - The smallest adverse response must be observed in one of the two field reference sediment treatments (Stations A1 or A3) with the lowest PCB concentrations.

– The highest sediment PCB concentration must yield either the largest adverse response or a response that is not statistically different from the largest adverse response.

- Amphibians – No revisions to the amphibian MATC are necessary. The MATC derivation considered only endpoints with a dose-response curve that could be fit using the probit model, linear regression, or Trimmed Spearman-Karber method, as detailed in Table 4.E.1 - Summary of Threshold Effects Concentrations for the Wood Frog Study, as Modeled with Vernal Pool Sediment tPCBs and Spatially Weighted Mean tPCBs, and Section E.4.3 – Concentration-Response Analysis – Toxicity Endpoints. Additional text will be added to the ERA to increase transparency of the statistical analyses.

- Fish – The endpoints evaluated in the fish toxicity studies have a number of sources of variability and uncertainty (e.g., test organism sensitivity, the mechanism by which PCBs and other COCs exert toxicity, and uncertainty about exposure concentrations). Data of this type cannot be reasonably expected to exhibit a monotonic dose-response shape. EPA will consider all available data and clearly identify the trials that exhibited the following:

  - Strong dose-response (using decision rules similar to those for benthic invertebrates).

  - Interrupted dose-response.

  - No dose-response (i.e., effects in highest concentration tested were not significantly greater than controls).

- Wildlife – For derivation of the mink and river otter MATC, the geometric mean of the NOAEL and LOAEL from the mink feeding study by Bursian et al. (2002) was used. For eagles, there are insufficient data to generate a dose-response curve; therefore, the toxicity threshold was used in establishing the MATC. EPA believes this is an acceptable approach for an endangered species such as bald eagle, rather than using a surrogate species for which a does-response relationship can be established.

- For all receptors, endpoints that do not demonstrate acceptable dose-response will be discussed qualitatively, as supporting evidence for the MATC and risk characterization.

6.E. Consider Only Sediment-Relevant Organisms

This recommendation for derivation of the MATCs applies to the benthic invertebrate endpoint only; therefore, the detailed response is provided in the response to General Issue 10.E.

6.F. Consider All Endpoints, Not Only the Most Sensitive

ERA agrees that the toxicity thresholds for multiple endpoints should be considered. However, the analysis must also take into consideration the sensitivity of the test endpoints, in terms of exposure type/duration, the tolerance of the test organisms, and the relevance of the measurement endpoint to the Assessment Endpoint. It should also be recognized that investigators commonly include collection of data for a variety of metrics that may not necessarily reflect the impact of a stressor on the test organism, and may therefore not be appropriate for development of a toxicity threshold. In the revised ERA, EPA will develop MATCs based on consideration of all species and endpoints relevant to the Assessment Endpoint, but with the recognition that the conclusions and actions based on the risk assessment must be protective.

For aquatic endpoints, EPA will, where possible based on adequate sample size, apply the species sensitivity distribution approach recommended by some Reviewers for this purpose. For both acute and chronic MATCs, the selected value will therefore reflect both an effect size and the percentage of species/endpoints exhibiting the specified level of effect.

The Reviewers' comments do not apply to fish and amphibians because the distribution of effects thresholds for these receptors was not truncated to include only the most sensitive endpoints. For wildlife, the MATCs are intended to be species-specific and thus a species sensitivity distribution approach would be inappropriate. For all receptors, endpoints that did not exhibit statistically significant responses (i.e., not used for MATC derivations) will be discussed in the revised ERA.

6.G. Truncate MATC at the Reference Concentration

EPA concurs with this recommendation, with the caveat that the reference concentration must represent a true reference condition, rather than contamination associated with GE.

For benthic invertebrates, Figure D.2-13 of the ERA presents the sediment tPCB concentrations at all toxicity test stations. At the reference stations (A1 and A3), all individual sediment samples except one had a tPCB concentration below 0.6 mg/kg. A single sediment sample of 5.4 mg/kg tPCB was observed at Station A3 in conjunction with the 7-d in situ study. The location of this sample was downstream of Dorothy Amos Park, a site contaminated with PCBs from the GE facility. Other localized elevated PCB concentrations (sediment, tissue) have been observed at or near this location. Based on the above criterion, the MATC for benthic invertebrates must exceed 0.6 mg/kg, but may fall below 5.4 mg/kg.

The Reviewer's comment would also apply to other receptors. For amphibians, the tissue MATC derived was greater than the reference concentration (with the exception of one anomalous metamorph tissue value), and the sediment MATC was above the reference sediment concentrations. The fish tissue MATCs (tPCB, TEQ) were also higher than the reference concentrations from Threemile Pond. Therefore, no modifications to MATC values are anticipated for these receptors.

The MATC of 2.65 mg/kg tPCBs in fish tissues is above the concentrations observed in all fish samples taken from the Upstream Reference Area (n=45) and from the Threemile Pond Reference Area (n=31) (Table 1.2-4 in current ERA). The sediment MATC for bald eagles was greater than the concentrations in reference sediment.

7. DEFINITION OF HQ AND RISK CATEGORIES

SUMMARY OF ISSUE:

Five Reviewers made a number of suggestions with regard to how risk estimates should be presented and described in the ERA. Some Reviewers did not understand how the hazard quotients (HQs) were being interpreted with regard to magnitude of risk. One Reviewer noted that the system used to categorize risk differed between aquatic and wildlife endpoints. In addition, some Reviewers suggested that it would be useful to estimate probabilities of exceeding MATCs or other benchmarks and to use cumulative frequency plots to show the relationship between exposure distributions and corresponding MATCs or other benchmarks. Reviewers noted that the use of value-laden terminology should be avoided.

RESPONSE:

As previously stated in EPA's response to Reviewers' questions prior to the Peer Review Meeting, EPA agrees that the terminology regarding HQs and MATCs in the ERA could benefit from further clarification as the term HQ was used in a number of different contexts. The approach described in the Risk Characterization section in the *Framework for Ecological Risk Assessment* (1998) guided what was done in the ERA. Considering this guidance, the HQ terminology was applied in the ERA for a number of different situations, as shown in Figure 9 and discussed below.

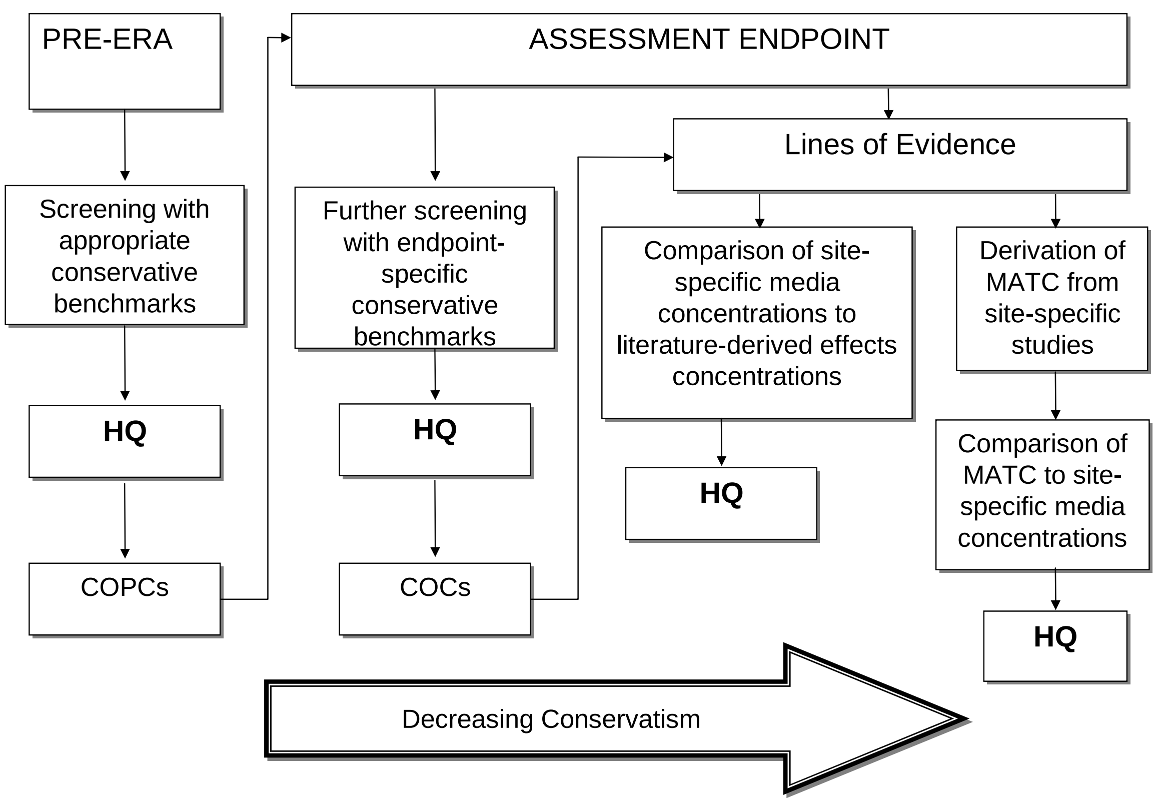


Figure 9 Use of Hazard Quotients in the ERA

Screening of COPCs

HQs were used as part of the Tier 1 screening of COPCs, as described in Appendix B to the ERA. In summary, HQs were calculated by dividing site-specific concentrations of individual COPCs by an appropriate conservative medium-specific benchmark concentration for the contaminant. The resulting HQs were grouped into ranges (e.g., 1≤HQ<10; 10≤HQ<100) for discussion purposes. In the Tier 1 evaluation, contaminants with an HQ < 1 were eliminated as COPCs. Contaminants with an HQ > 1 were retained for further consideration in Tiers II and III, which included comparisons to background and other screening criteria. No judgments regarding the risk associated with the magnitude of an HQ > 1 were stated or implied in this screening assessment.

After the pre-ERA, additional screening was performed for individual Assessment Endpoints to refine the list of COCs for the specific endpoint. HQs were calculated using conservative benchmarks, and COPCs with an HQ < 1 were eliminated from further consideration as a COC in the risk assessment for that endpoint. Again, no conclusions regarding the risk associated with the magnitude of an HQ > 1 were drawn from this screening evaluation.

Risk Descriptions in the ERA for Aquatic Assessment Endpoints

As discussed in Section D.4.2.1 of the ERA, HQs were used to quantify the degree to which chemistry measurements exceeded environmental benchmarks considered protective of benthic invertebrates. In theory, adverse ecological responses are possible if any HQ exceeds 1 (i.e., when the exposure exceeds the protective benchmark). However, most benchmarks were developed for screening purposes and the inherent conservatism in such benchmarks leads to a high percentage of false positives (Chapman and Mann 1999). To address this uncertainty, the HQ assessment for the aquatic endpoints considered multiple benchmarks, resulting in a range of calculated HQs. For each contaminant and medium, the full range of calculated HQs was considered and an HQ was also calculated using the median of the applicable benchmarks. This median value represented an intermediate level of protection, and the extremes of the HQ distribution represented the upper and lower bounds.

These calculated HQs were considered to be one line of evidence, which was then compared with two additional lines (results of toxicity tests and field studies with semi-qualitative observations). Within the PSA, all of these lines of information were used to derive an assessment of risk in the ERA, termed low, intermediate, or high. The term "no risk" should have been used only in the situation where the exposure was not occurring at a level exceeding the most conservative benchmark values; in most cases contaminants in that category were screened out in the Pre-ERA.

The site-specific toxicity studies provided the data used to calculate the MATC. This calculation was based upon selecting the concentration at which an effect of biological concern was observed, and no uncertainty factors were applied in the calculation. In this case, HQs represent a comparison of the MATC to PCB concentrations in the appropriate media. An HQ > 1 was considered to be an indication of risk ("risk of harm"), as concentrations in site media exceed those observed in site media used in the toxicity tests which resulted in an effect of biological concern. An HQ of 10 was not interpreted in the ERA as being ten times worse than an HQ of 1.

In one location in the ERA, there is a statement saying that for fish most of the HQs are below 3 as a point of observation only.

Risk Descriptions in the ERA for Wildlife Endpoints

The wildlife endpoints did not always benefit from having all lines of evidence; in these circumstances, the risk descriptions were more uncertain. For the modeled line of evidence, which expresses risk as a function of probability of exposure and effect concentrations, there was also the distinction between situations where a dose-response curve was available versus those where data from the literature limited it to a range of effects thresholds. In the case of these numerical comparisons, the risk criteria used in the ecological risk assessment for the Calcasieu Estuary (EPA 2002) were used. The rationale for the risk criteria were provided on pages 6-27 and 6-28 in Section 6 of the ERA. The following criteria were used to describe risk for the wildlife endpoints:

Scenarios with effects data for the representative species (or a reasonable surrogate) that are not threatened or endangered:

- If the probability of 10% or greater effect (or of exceeding the NOAEL) was less than 20%, then the risk was categorized as low (Figure 6.7-1).
- If the probability of 20% or greater effect (or of exceeding the LOAEL) was greater than 50%, then the risk was categorized as high (Figure 6.7-1).
- All other outcomes were categorized as intermediate risk (Figure 6.7-1).

Scenarios with effects data for the representative threatened and endangered species (or a reasonable surrogate):

- If the probability of 10% or greater effect (or of exceeding the NOAEL) was less than 20%, then the risk was categorized as low.
- If the probability of 10% or greater effect (or of exceeding the NOAEL) was greater than 50%, then the risk was categorized as high.
- All other outcomes were categorized as intermediate risk.

Scenarios with threshold ranges derived for the representative species (both those that are threatened or endangered and those that are not):

- If the probability of exceeding the threshold for the most sensitive species was less than 20%, the risk was categorized as low.
- If the probability of exceeding the threshold for the most tolerant species was greater than 20%, the risk was categorized as high.
- All other outcomes were categorized as intermediate risk.

The above risk criteria were used for all representative wildlife species, including threatened and endangered species, in the current ERA. The risk criteria will not be changed in the revised ERA. The risk estimates derived using the criteria described above were combined with the other available lines of evidence and associated weights to derive a final characterization as low, medium, or high risk. The details are described in the Weight-of-Evidence section under Risk Characterization in each appendix.

Risk Descriptions Below the PSA

Risk estimates generated by applying the MATCs developed using data collected in the PSA downstream may be more uncertain due to the potential differences between the areas in habitat, species, etc. Therefore, more subjective terminology was used in an attempt to further qualify these risk estimates. However, EPA will revisit the discussion of downstream risks to improve the distinction between the estimation of the likelihood of risk, and the uncertainty surrounding the estimate.

Revisions to ERA

EPA agrees that the use of HQs in the ERA would benefit from some clarification and will implement the following changes in the revised ERA:

- Provide cumulative frequency plots for exposure of aquatic endpoints (as was done for wildlife), and estimate probabilities of exceeding MATCs or other benchmarks.
- Develop criteria to classify risk to aquatic endpoints that are consistent with the existing criteria used to classify risk for wildlife endpoints.
- Avoid use of value-laden or undefined terms to describe risk.

8. CONSISTENCY OF WOE ANALYSES

SUMMARY OF ISSUE:

All seven Reviewers had general or Assessment Endpoint-specific comments regarding the weight-of-evidence (WOE) approach that was incorporated in the ERA.

- While several Reviewers disagreed with a weighting assigned to a particular measurement endpoints, the majority of comments centered on issues associated with clarity and transparency when applying the Massachusetts Weight-of-Evidence Approach to the lines of evidence available for this assessment.
- One Reviewer recommended that the advantages and disadvantages of the different lines of evidence be discussed and that lack of concordance between lines of evidence be explained.
- Most of the Reviewers commented that some of the GE field studies had been weighted too highly in the WOE evaluation. Some Reviewers also commented that the GE, and in

some cases, EPA field studies should be weighted more highly in comparison to the modeled lines of evidence. Otherwise, Reviewers were generally satisfied with the weighting assigned to the EPA field studies.

- Six Reviewers commented specifically on the weight assigned to information from fish field studies (both GE and EPA). Four Reviewers noted that the weighting assigned to this line of evidence appeared low. One Reviewer presented a dissenting opinion that "the literature data and the toxicity studies should be assigned more weight" in the ERA. One Reviewer commented that potential risk to fish is underestimated, based on the observation that the presence of reproducing fish (shown in field studies) is insufficient for a determination of a healthy population.

RESPONSE:

8.A. Weight-of-Evidence Approach

EPA agrees that the clarity and transparency in the presentation of the weight-of-evidence approach (Menzie et al. 1996) can be improved in the ERA. Specifically, Reviewers had comments regarding the level of transparency provided when applying this methodology to the measurement endpoints evaluated. The most frequent observation was regarding the level of detail provided to justify the weight, and one Reviewer questioned the consistency when assigning weights to specific attributes and the resulting score to each measurement endpoint.

The weight-of-evidence approach used in the ERA follows the approach described in the *Massachusetts Weight-of-Evidence Special Report* (MDEP) (Menzie et al. 1996) and is described in Section 2.9 of the document. Although most Reviewers commented favorably on the approach or offered no comment, there were aspects of the approach that caused confusion. In particular, the Reviewers found that some of the possible options in the weight-of-evidence summary tables were difficult to interpret. For example, "a combination of 'no evidence' for a 'low magnitude' of harm … could be interpreted that the study would have been able to detect a low magnitude of harm but didn't or that the study did not find evidence for harm but there nevertheless could have been low magnitude harm that was not picked up." The option of 'no evidence of harm' and 'high magnitude of harm' caused some confusion as the option appears contradictory.

To address these comments, EPA will expand the general weight-of-evidence discussion in Section 2.0 to include a detailed presentation of the weighting process. In addition, EPA will review the weight-of-evidence discussion for each Assessment Endpoint on a section-by-section basis. Emphasis will be placed on providing a consistent level of detail when evaluating all lines of evidence and confirming that all weight-of-evidence tables are formatted in a consistent fashion. Where applicable, professional judgment will be presented in a manner that better allows the reader to understand the thought process associated with these determinations.

8.B. Advantages/Disadvantages of Lines of Evidence/Lack of Concordance

As described in Section 2.9 of the ERA, there are three general lines of evidence under which most measurement endpoints fall:

- Field survey data that provide a characterization of the animals at the site.
- Toxicity testing that indicates whether the contaminated media are toxic to representative species.
- Effects data from the literature compared to contaminant concentrations measured in site media.

Conclusions regarding risk have a high degree of confidence when multiple lines of evidence suggest a similar conclusion regarding risk. When the lines of evidence are not in agreement, however, the risk assessor needs to consider the possible explanations for the lack of concordance between the lines of evidence.

For example, a comparison of site-specific estimates of exposure to concentrations reported to be toxic in the literature suggests that the receptor is at high risk, yet field surveys may show that the receptor is present in both contaminated and uncontaminated areas at similar densities. There are several possible explanations for these observations, some of which are associated with risk while others are not. Some of these explanations include:

- The species used in the test reported in the literature is more sensitive than the species of concern at the site (no risk to the representative species, but possible risk to other species).
- Receptors in the contaminated area are being adversely affected but are being replaced by immigrants from nearby uncontaminated areas (risk).
- The individuals comprising the local population have adapted and become tolerant to the contaminant (risk).
- The field survey was not designed to or lacked the statistical power to detect differences in receptor abundance between uncontaminated and contaminated areas (undetermined).

Thus, it is possible to have a lack of concurrence between lines of evidence, yet reach a conclusion regarding risk, but with more uncertainty associated with this conclusion.

8.C. Weighting of Field Studies

EPA will reconsider the WOE designation for the field studies in the revised ERA based on the Reviewers' recommendations. The fish field studies received considerable attention in the Reviewers' comments. The fish field studies were given a low/moderate WOE designation in Table 5.4-4 because they were not designed to detect any effects attributable across a concentration gradient of PCBs, or to detect effects other than at a gross population level, and did not include an evaluation of reference areas. However, the studies did demonstrate that there is a viable fish community present in the river. Had the information collected in the field studies not been given significant consideration, the overall risk conclusion would have been closer to the moderate/high designation indicated by the other lines of evidence. Therefore, it is clear that the transparency in the WOE evaluation needs improvement.

As noted by one Reviewer, the presence of self-sustaining fish populations does not equate to a determination of population health. However, the presence of largemouth bass and other fish species provides evidence that effects such as larval deformities have not translated to readily observable impacts using the field survey methods.

EPA will revise the ERA to provide a more transparent assessment of how field studies are used in the risk characterization.

9. ECOLOGICAL LEVEL OF ORGANIZATION

SUMMARY OF ISSUE:

Five Reviewers provided comments about the "level of ecological organization (i.e., individual-, population-, community-, ecosystem-level) for which protection of the Assessment Endpoint is sought." Some Reviewers thought that "long-term persistence of populations of target taxa" is the goal of the assessment. Other Reviewers did not offer an opinion on the appropriate level of organization, but did request that EPA state the protection goal for each Assessment Endpoint.

RESPONSE:

The guidance for ecological risk assessments under Superfund (EPA 1997) states that "appropriate Assessment Endpoints differ from site to site, and can be at one or more levels of biological organization." Further, "appropriate Assessment Endpoints might involve local populations of a particular species, community-level integrity and/or habitat preservation."

It is also stated in the guidance (EPA 1997), "it may not be practical or technically possible to document existing ecological impacts [i.e., population-level impacts], either due to limited technique resolution, the localized nature of the actual impact, or limitations resulting from biological or ecological constraints of the field measurements." As a result, impacts at lower levels of organization (e.g., adverse effects on survival of individuals) are often used to infer possible impacts at higher levels of organization (e.g., persistence of local populations) (EPA 1997).

In the Ecological Risk Assessment and Risk Management Principles (EPA 1999), it is stated that "Ecological risk assessments incorporate a wide range of tests and studies to either directly estimate community effects . . . or indirectly predict local population-level effects . . ., both of which can contribute to estimating ecological risk. Superfund remedial actions generally should not be designed to protect organisms on an individual basis . . ., but to protect local populations and communities of biota. Levels that are expected to protect local populations and communities can be estimated by extrapolating from effects on individuals and groups of individuals using a lines-of-evidence approach."

Following this guidance, it is stated on page 2-66 of the ERA, "Although many of the [assessment and measurement] endpoints presented are linked to organism-level effects (e.g., survival and reproduction), these endpoints are expected to be strong indicators of potential local population-level effects . . ." Thus, the general focus of the ERA on measuring effects using established techniques at the organism level and extrapolating the results, using accepted methods, to the local population is appropriate.

The following guidelines were used in the ERA in determining at what level of organization risks should be estimated:

- If risk estimates at the individual level were low (e.g., tree swallows), then risk analyses at higher levels of organization were not performed because hierarchy theory states that effects at the population or higher levels of organization cannot be occurring in the absence of organism-level effects.

- If risk estimates at the individual level were high and such estimates were consistent across multiple lines of evidence (e.g., mink), then risk analyses at the higher levels of organization were not performed because EPA believed that the answer was obvious (i.e., local populations are being adversely impacted) and thus there was little benefit to performing the analyses.

- If risk estimates did not fall into either category and methods were readily available to estimate risks at higher levels of organization (e.g., amphibians, benthic invertebrates), then EPA proceeded to conduct further analyses. The purpose of such analyses was to improve the understanding of the potential ecological consequences of impacts that may be occurring at the organism level of organization. Where methods were unavailable or information was lacking to conduct risk analyses at higher levels of organization (e.g., many of the bird receptors), no further analyses were undertaken.

- For threatened and endangered species, no risk analyses were undertaken above the individual level of organization. As stated on pages 6-28, "Any effect to threatened and endangered species is cause for concern (Massachusetts Office of Environmental Affairs 1999; Massachusetts Division of Fisheries and Wildlife 2003; United States Congress 1973, Endangered Species Act)". Thus, when risks are intermediate or high at the individual level of organization, there is no need to pursue analyses at higher levels of organization.

While these guidelines were used in the development of the ERA, they were not explicitly stated. EPA will present the guidelines in the revised ERA and will provide discussion of the relevance of effects at the organism and lower levels of organization to understanding impacts at the local population level of organization.

10. USE OF UCL POINT ESTIMATE FOR CONCENTRATION TERM IN MONTE CARLO ANALYSIS

SUMMARY OF ISSUE:

Three Reviewers commented on the use of the upper 95% confidence limit (UCL) on the mean as the point estimate for chemical concentration variables in the wildlife exposure analyses. One Reviewer, in particular, felt that "variability and uncertainty associated with PCB/TEQ concentrations are not represented" in the current exposure analyses. This Reviewer suggested that the "Monte Carlo analyses should provide what is believed to be the reasonable most-likely exposure distribution based on expected distributions of all [input] parameters." Although these Reviewers commented regarding the use of the UCL to describe prey concentration variables, there was no consensus on this point. One Reviewer stated that the methodology and rationale developed for calculating exposure concentrations in the ERA was "excellent."

RESPONSE:

The approach for calculating exposure point concentrations for the prey concentration variables in the wildlife exposure analyses was provided in Section 6.3 and Appendix C.5 in the ERA. These sections also provided the rationale for why the UCL was used to describe prey concentration variables in the Monte Carlo analyses. A summary of what was discussed in Section 6.3 and Appendix C.5 is provided below:

- Concentrations of COCs vary spatially and temporally in prey. The representative wildlife species forage over distances ranging from tens of meters to greater than 10 km. Thus, individuals tend to integrate spatial variation in the tissue concentrations in their prey over time. Therefore, estimates of the central tendency (i.e., arithmetic means) are used in the exposure model as an expression of the spatial and temporal averaging of concentrations of COCs in prey tissues (EPA 1999). In the Monte Carlo analysis, it was assumed that the spatially and temporally averaged exposure estimate did not vary between individuals foraging in the same area. Thus, the point estimate of centrality was the minimum of (1) the 95% UCL, calculated using the Land H-statistic (assuming data are lognormally distributed), or (2) the maximum concentration measured.

- EPA (1992) states that "because of the uncertainty associated with estimating the true average concentration at a site, the 95 percent upper confidence limit (UCL) of the arithmetic mean should be used for this variable." For lognormal data, EPA (1992) recommends the Land method using the H-statistic. Several authors (e.g., Ott 1995; Seiler and Alvarez 1996; Hattis and Burmaster 1994) have argued that concentrations of contaminants in environmental media tend to be lognormally distributed and that this may be expected because of mechanistic reasons. Current EPA guidance (EPA 1997; also see Haimes et al. 1994) states that distributions should be chosen for input variables on the basis of mechanistic or theoretical reasons, if possible, because such distributions have the highest degree of confidence. As a result, concentrations of contaminants in prey were assumed to be lognormally distributed in this ERA, and hence the Land H-statistic was used to estimate the 95% UCL. That said, it is recognized that the Land method can produce high values for the UCL, particularly when data are not lognormally distributed, samples size is small, or variation is high (Singh et al. 1997; Schultz and Griffin 1999). EPA (1992; 2001, Appendix C) guidance recognized this problem and recommended that the maximum detected concentration be used when the calculated UCL exceeds this value. For a discussion of the results of testing for lognormality of prey concentration data sets, please refer to Section 6.3 of the current ERA. This guidance was followed in the ERA.

The UCL accounts for sampling uncertainty regarding the true average concentration at a site. As sample size increases, the difference between the UCL and the true average concentration decreases. Thus, use of the UCL accounts for uncertainty introduced by the limitations of sampling (i.e., the inability to sample the entire population). It is for this reason that EPA (1992, 2001) recommends that UCLs be used for prey concentration variables in probabilistic exposure analyses that rely on Monte Carlo methods to propagate uncertainties. As noted above, this guidance was followed in this assessment.

11. ASSESSMENT ENDPOINT—BENTHIC INVERTEBRATES

SUMMARY OF ISSUE:

Reviewers had a number of comments on the benthic invertebrate Assessment Endpoint. These comments are presented below, grouped into a number of general categories. Each of these bulleted categories is treated separately in the responses that follow.

- Some Reviewers recommended supplemental statistical analyses, specifically with regard to analysis of the benthic invertebrate community data: Three Reviewers recommended incorporation of diversity indices, and indicated the Shannon-Wiener index (presented to Reviewers in December 2003 as part of responses to initial questions) might not be the optimal measure for benthic communities in which a few species dominate. Two Reviewers indicated that the increased use of power analysis should be applied to analysis of the benthic data. This was related to the comment that both effect size and statistical significance are important.

  One Reviewer indicated that there are several additional techniques that could be applied to the benthic invertebrate effects data. These include a log-linear categorical analysis and canonical correlation analysis.

  Six Reviewers discussed the quantitative evaluations of measurement endpoints (particularly for benthic community structure) that were performed by both EPA and GE in response to initial Reviewer questions. The Reviewers requested that the evaluations be incorporated in the ERA and considered in the WOE evaluation.

  Three Reviewers indicated that, based of the multiple regression analyses conducted on behalf of GE, PCBs appear to explain a small fraction of the total variability of benthic communities. Accordingly, it was recommended that EPA consider both effect size and statistical significance in the community evaluations.

- Three Reviewers commented that the sediment concentration data were highly variable over very small spatial scales, and that this variability in sediment concentration data created difficulty in the identification of dose-response relationships for toxicity endpoints. The Reviewers differed in opinion over the treatment of the variability in PCB concentrations in the statistical analyses. One Reviewer stated that EPA generally did a good job of describing the variability and attempting to account for it, whereas another Reviewer suggested that the variability in the tPCB data needs to be assessed to determine if the differences observed are statistically different between the test stations. It is unclear whether this comment refers to the measures of central tendency or the variations among sampling locations.

- Four Reviewers commented on the need to provide an improved discussion of the linkage between the results of the benthic community surveys and the toxicity tests. In particular, the Reviewers encouraged EPA to articulate more clearly when there are differences between the results of the studies and to provide discussion of chemistry differences between sampling programs.

- Four Reviewers commented that the abundance and distribution of benthic macroinvertebrates in Housatonic River sediment varies considerably among samples, creating difficulty in separation of the effects of tPCB concentrations from the effects of other variables. Reviewers requested that a more complete discussion of the role of substrate on the relationship between PCB exposure and biological effects be provided.

- Some Reviewers commented that the benthic invertebrate endpoint receptor group was either not adequately defined, or that some measurement endpoints were inappropriate for the group. There were several related comments, including:

  - One Reviewer indicated that risks to receptors via the epifauna exposure pathway are not considered in the ERA.

  - Three Reviewers indicated that cladocerans are inappropriate test species for representing risks to benthic infauna.

  - One Reviewer indicated that risk conclusions for benthic invertebrates would be applicable only for soft sediment where benthic infauna are found.

  - One Reviewer indicated that the benthic invertebrate ERA did not address worst-case risks, which would require assessment of both infauna and epifauna.

- Three Reviewers commented that the sampling protocol for benthic tissue at downstream stations differed somewhat from upstream stations, and that this may have implications for the correspondence between tissue and sediment chemistry results. The Reviewers commented that this needs to be made clear in the report and described in the uncertainty section.

  One Reviewer commented on an apparent discrepancy between the sediment collection methods described in the SIWP and the collections actually performed in conjunction with the benthic community grab sampling program. Specifically, the sediment sampled in the upper portions of the PSA (Reach 5A) appeared not to be “soft and depositional.”

- One Reviewer commented that a complete discussion of communities, substrates, riverine conditions, and organic carbon is needed to determine if there is a relationship between tPCBs and toxicity.

RESPONSE:

11.A. Additional Analyses

Diversity Indices

EPA will retain the Shannon-Wiener evaluation, as specifically requested by one Reviewer. However, EPA will also present the Simpson's Index, as requested by three Reviewers. Both the effect size and statistical significance of the indices will be explored and reported.

Power Analysis

Some Reviewers indicated that power analyses should be incorporated in the ERA. The utility of power analyses depends on the context in which they are being considered. The following approaches will be applied in the revised ERA:

- Situations when power analysis is appropriate – Where inferential statistics were applied, and single parameters are considered, power analyses are useful additions to the ERA. For example, the invertebrate toxicity tests will be assigned Minimum Significant Differences (MSDs) such that the Reviewers can assess the ability of the tests to detect various effects sizes.

- Situations where power analysis is inappropriate – Where statistics are descriptive only (i.e., non-inferential), power analyses are not required. For example, power analyses would not be applied to cluster analysis results. Also, power analyses are not appropriate for situations in which a statistically significant effect was observed, since there is no question regarding a Type II error in these cases.

- Situations where power analysis may be useful – In some cases, power analysis may be applicable for multivariate methods, but is not straightforward, and would require careful analysis. EPA will evaluate these cases and determine whether power analysis would provide additional useful information to the revised ERA.

Additional Multivariate Methods

EPA applied multiple multivariate methods to evaluate the benthic community data, including multidimensional scaling (MDS), cluster analysis, and relative rank plot approaches. These methods exhibited strong concordance in findings. Additional statistical evaluations will be incorporated in the revised ERA, and EPA will specifically consider the use of canonical correlation analysis to explore relationships between data sets.

Other Data-Related Recommendations

EPA agrees that the supplemental analyses performed as a response to Reviewers' questions will improve the evaluation of the benthic community data, and will integrate the findings into the revised ERA.

Specific additions to the benthic invertebrate component of the ERA will include:

- Regression analyses conducted by both GE and EPA relating benthic indicators to PCBs and substrate parameters such as percent fines and TOC.

- Plots of indicator taxa as a function of PCB concentration and substrate variables.

- Calculation of additional benthic metrics, such as the Simpson's Diversity Index, and relation to PCB concentration and substrate.

For each analysis, the information will be reviewed in terms of uncertainties, limitations, and findings (or lack thereof) of ecological harm.

EPA agrees that effect sizes and contribution to total variability should be considered in the ERA. However, EPA also believes that sources of natural variation, measurement and analytical variability, and appropriateness of the statistical models are also important considerations for interpreting the results. EPA will include more discussion of this issue in the revised ERA, in conjunction with existing statistical methods and the additional methods described in 5.A.

11.B. Variable Sediment PCB Concentrations and Implications for Dose-Response Assessment

EPA agrees that sediment PCB concentrations were highly variable over very small spatial scales, and EPA performed multiple analyses to evaluate this situation and presented these in the ERA. One consequence of the variability in the PCB concentration data is that analyses have the potential to be sensitive to the selection of the exposure data matched to effects. Five Reviewers provided comments on this issue (please refer to the response to General Issue 7). EPA will amend the dose-response analyses as requested, and will provide additional information on the spatial relation of samples collected during different sampling events. EPA notes that because sediment PCB variability is high, use of exposure concentrations based on a single ("most synoptic") sample as suggested by Reviewers also is subject to uncertainty.

EPA does not agree that the variability in sediment concentration data created difficulty in the identification of dose-response relationships for toxicity endpoints. The uncertainty that resulted from the PCB variability mentioned above was significant; however, it did not impede the development of dose-response curves. Although variability in the independent variable reduced the precision in the dose-response curves, the overall relationship between PCB and toxicity is apparent in spite of this variability. This is true whether the "median" or "most synoptic" data are chosen. Using the guidelines for establishing "dose-response" described in the response to General Issue 7, the majority of toxicity endpoints exhibited acceptable dose-response relationships.

EPA does not agree that the statistical significance of pairwise comparisons in PCB concentrations among sites should affect the dose-response analysis approach. Although the within-station variability affects statistical power of the dose-response analysis, the validity of a model that identifies a significant dose-response cannot be questioned simply on the basis of variability within or among individual stations. In the revised ERA, both central tendencies and variances will be compared across sampling locations (ANOVAs and pairwise comparisons), as supporting information.

11.C. Dose-Response Relationships

EPA concurs with the Reviewer comments, and will describe more clearly the similarities, differences, and uncertainties for these components of the sediment quality triad. The characterization of risk at Stations 7 and 8, in particular, will benefit from a more detailed discussion of dose-response. Application of the sediment quality triad at this site is somewhat more complicated than at other sites due to the multiple measurements of sediment chemistry made for each triad component, combined with the high small-scale variability in sediment PCB concentrations.

Although EPA recognizes that variability in sediment parameters made this analysis more complicated, EPA believes this subject was adequately discussed in Section D.2 of the ERA. ERA Section D.2.4.4.3 and Figure D.2-13 in the ERA address the differences in PCB exposure across Triad components for Stations 7 and 8. In the risk characterization (ERA Section D.4.6.3, lines 13-19) there is further discussion regarding the lack of intra-station consistency in PCB exposure concentrations, which complicated the station-by-station risk characterization. In the revised ERA, EPA will advance the discussion by explicitly evaluating the consistency of biological responses, in terms of both consistency of results across Triad components and consistency of results across sampling locations.

11.D. Variation in Benthic Assemblages

EPA agrees with the Reviewers that the abundance and distribution of benthic macroinvertebrates in Housatonic River sediment varied considerably among samples. This was minimized by partitioning the sample stations into two broad substrate types. However, it was not possible to control for all micro-habitat variables; therefore, some of these factors still may contribute to the variability which was observed. Examples of these factors include:

- The distribution of benthic macroinvertebrates in the environment is typically not uniform or random, but clumped. Therefore, the abundance of organisms in samples taken at random is inherently expected to contain variable numbers of organisms.

- Chironomids and other organisms display clear preferences in the types of substrates that they frequent. For example, *Cladotanytarsus* and *Cryptochironomus* spp. were collected commonly at sandy sites, but much less so at fines-dominated sites. The opposite is true for the *Tanytarsus guerlus* grp. and *Einfeldia* spp.

- The amount of TOC present varies among sites with similar concentrations of fines.

In summary, significant variability in benthic community samples is expected. In the revised ERA, EPA will provide additional discussion of the influence of substrate on benthic assemblages. However, natural variations will persist that cannot be addressed through the application of increasingly sophisticated and reductionist mathematical models.

11.E. Representative Test Species and Implications for Risk Characterization

EPA does not agree with the comment that epifauna were not included in the definition of "benthic invertebrates." In EPA's response to the request in the preliminary questions from the Reviewers to define "benthic invertebrates," the following definition was provided:

> ". . . benthic invertebrates refer to invertebrates that are exposed via sediment and porewater uptake pathways. These organisms include "infauna" and "epifauna" but exclude "water column invertebrates."

EPA's definition of epifauna in the ERA includes organisms that obtain significant exposure from both the sediment bed (particulates and pore water) and the water column (suspended particulates and water column). The definition excludes only those organisms that feed exclusively from water column sources (e.g., zooplankton, copepods, rotifers, etc.).

The measurement endpoints for the benthic invertebrate ERA clearly indicate that epifauna were considered in the ERA. Both species tested in the laboratory toxicity study (*Hyalella* and *Chironomus*) are epifaunal taxa. Chironomids also represented a significant portion of the field-collected communities and were studied as indicator taxa. The benthic grab samples contained both epifauna and infauna, and epifaunal insect larvae dominated the abundance and biomass of most samples.

EPA does not agree that cladocerans are inappropriate test species for representing risks to benthic invertebrates. Toxicity to cladocerans (i.e., *Daphnia* and *Ceriodaphnia*) has relevance to effects on epifauna and is therefore appropriate for evaluating sediment toxicity.

- EPA (2000a) provides guidance for testing sediment toxicity to freshwater invertebrates, and specifically mentions cladocerans as being among the "standard methods" for assessing sediment toxicity (Section 1.1.4):

  > "A variety of standard methods have been developed for assessing the toxicity of contaminants associated with sediment using amphipods, midges, polychaetes, oligochaetes, mayflies, or *cladocerans* (i.e., ASTM 1999a; ASTM 1999b; ASTM 1999c; ASTM 1999d; EPA 1994a; EPA 1994b; Environment Canada 1997a; Environment Canada 1997b)." (*emphasis added*)

Information contained in the protocol documents referenced by EPA (2000a) (particularly ASTM testing manuals) supports the use of daphnid testing as a "standard method" for sediment toxicity. ASTM (2001; Volume 11.05; Designation E 1706 – 00) contains a description of standard methods for measuring the toxicity of sediment-associated contaminants with freshwater invertebrates, and Annex A2 of the ASTM (2001) document contains specific guidance for conducting sediment toxicity tests with both *Daphnia magna* and *Ceriodaphnia dubia*.

EPA agrees that the linkage between cladoceran toxicity and infauna risks may be indirect; however, the cladoceran tests were designed to indicate risks to epifauna. The sensitivity of cladoceran taxa to the sediment-based exposure pathway has been documented not only in this study (in situ toxicity by Wright State University; EVS 2003), but also in other applications (Burton et al. 1989). While cladoceran life history is somewhat more dependent on water-column exposures relative to the other epifauna test organisms (i.e., *Hyalella, Chironomus*), they are still considered useful indicators of sediment-mediated toxicity.

To address the comment, EPA will include a more specific and detailed definition of the "benthic invertebrate" receptor group.

EPA does not agree with some of the Reviewer comments related to risk conclusions, because they are based on an incorrect assumption regarding the inclusion/exclusion of epifauna, in both the receptor definition and the applied measurement endpoints:

- Risk conclusions for benthic invertebrates are applicable only to soft sediment – The Work Plan for the benthic invertebrate sampling program discussed sampling soft bottoms. This was an operational definition, in the sense that no collections were attempted from rock and cobble bottoms that could not be penetrated by a grab sampler. The substrates sampled included both silt/clays and sands; there is very little rock/cobble bottom in the PSA. Because the sampling covered a wide and representative range of sediment TOC and particle size and the geographic extent of the PSA, and explicitly considered epifauna in both finer-grained and coarser-grained substrates, EPA believes the data on benthic communities reflect the range of conditions in the PSA.

- Worst-case risks not addressed – The Reviewer indicated that a worst-case assessment would require assessment of both infauna and epifauna. As discussed above, EPA did consider both infauna and epifauna. Although not every epifaunal taxon was assessed via toxicity tests, the range in organism life history and sensitivity was sufficient to make conclusions regarding risks to other taxa. Furthermore, the community assessment considered all taxa associated with the sediment bed, including epifauna.

11.F. Field Sampling

EPA agrees that more detailed discussion of the potential for sampling differences (upstream versus downstream) is appropriate, and will include discussion in the methods and uncertainties sections of the revised ERA. EPA also agrees that the physical separation between the D-net areas and the sediment chemistry grabs may be sufficiently large that the tissue chemistry and sediment chemistry results cannot be considered truly synoptic, particularly considering the small-scale variability in sediment PCB concentrations. This comment applies to both upstream and downstream locations. The tissue concentrations in the D-net samples cannot be precisely linked to sediment chemistry data, because the material that passed through the D-net was not submitted for chemical analysis.

EPA agrees that the samples in Reach 5A were coarser-grained and lower in TOC relative to the downstream reaches, but notes that the sandy substrates that predominate in the upstream areas are still "soft" in comparison with rock and cobble substrates, which could not be effectively sampled with a Ponar grab sampler. It is this relative definition of "soft" that was used in the work plan for the benthic invertebrate study. In the implementation of the SIWP, EPA targeted sediment that was both: (a) representative of the overall substrate condition in the sampled reaches; and (b) similar to reference locations, to facilitate statistical comparisons. The particle size and TOC data indicate that both objectives were satisfied.

Within Reach 5A, any truly depositional (cohesive) substrate is rare; overall the substrate is sandy and low-organic-carbon sediment. Although there are some small areas of very coarse material within Reach 5A (i.e., cobble/riffle habitat), the sandy substrate emphasized in the field collections is predominant in the reach. Sampling of small pockets of finer material would have been difficult to achieve in Reach 5A, and would have resulted in a biased sampling of Reach 5A substrate and associated biological communities. In the revised ERA, EPA will provide a more detailed discussion of choice of substrate during the benthic community sampling program.

11.G. Relationship to PCB

EPA does not agree with the statement by a Reviewer that the relationship to tPCBs and toxicity is tenuous. Among sites with relatively similar ranges of TOC and fines in the substrate, the abundance of indicator taxa appears to be strongly related to the tPCB concentration, although some species appear to be more tolerant of elevated tPCB concentrations than others. To address this Reviewer's comment, EPA will provide a more comprehensive evaluation of the available habitat information and the linkage between these measures and dose-response for PCBs in the revised ERA.

12. ASSESSMENT ENDPOINT - AMPHIBIANS

SUMMARY OF ISSUE:

Five Reviewers commented on EPA's leopard frog study. All five Reviewers remarked on the use of leopard frogs obtained from a biological supply company as reference frogs due to the lack of indigenous leopard frogs at the designated reference area. One Reviewer stated that comparisons to the purchased reference frogs were not valid; however, two Reviewers disagreed with that position. Two Reviewers questioned the small sample size used in the study, and one Reviewer questioned whether the timing of the field collections could have led to the conclusion of impact on leopard frog oocyte development due to PCBs in the PSA.

Five Reviewers commented on the wood frog population modeling, most requesting that a better explanation of the selection of model parameters be provided. Two Reviewers specifically requested that a sensitivity analysis of model parameters be conducted. Three Reviewers indicated that there should be more discussion of the model results in the context of the field studies and one recommended that the results of the GE-funded wood frog study (Resetarits, 2002) be incorporated into the modeling projections.

RESPONSE:

12.A. Leopard Frog Study

The discussion of the leopard frog study in the revised ERA will be expanded to clarify why, in the absence of viable frogs collected from reference areas, EPA believes frogs from Carolina Biological Supply (CBS) provided an appropriate control for analysis of potential effects in uncontaminated media. The discussion will explain why EPA believes that the use of the control frogs is more robust than using data from the literature, particularly since the leopard frogs used in the laboratory cross-over study spent several months (106 days) in the reference site (Muddy Pond) sediment/water. Additional discussion will also be included regarding the fertilization of the control frog egg masses, and the subsequent laboratory culture of the larvae through metamorphosis in Muddy Pond media.

EPA carefully considered sample size, visual interpretation of data, variability of data, and the uncertainty associated with findings in evaluating the leopard frog data, as discussed in Appendix E, Section 3.2.1 of the ERA. Inferential statistics for dose-response relationships were not considered appropriate for the analysis of leopard frogs effects data because:

- The magnitude and range of responses observed were not well suited to the detection of a dose-response relationship; the leopard frogs exhibited a threshold response for various endpoints. Threshold responses are not amenable to the use of traditional regression-based approaches; therefore, such data distributions are better evaluated via visual interpretation (figures) and examination of average response data (for a given endpoint) at each station.

- Sample sizes were small, primarily as a result of low availability of test animals from the field. A sample size of 5 was common with these data, and although correlation analyses can be conducted, the relationship must be very strong to be considered statistically significant. This is an unreasonable expectation for biological response data.

- A limited number of statistical tests could be conducted. Because fertilization of the egg masses from field-collected females was unsuccessful, there was no biological relationship between female PCB body burdens or reproductive tissue endpoints (i.e., percent Stage VI oocytes, percent malformed sperm cells) and larval/developmental effects endpoints (i.e., incidence of metamorphosis).

  EPA recognizes that visual interpretation alone of a small data set may lead to errors in conclusions. Therefore, the visual inspection of the juvenile effects data was considered together with other lines of evidence in determining conclusions of potential risk to leopard frogs:

  – The adult leopard frog specimens collected from the contaminated pools in the PSA showed signs of reproductive stress, and sample sizes were considerably higher than those associated with the juvenile effects endpoints. The timing of the collection of the animals coincided with the normal onset of reproductive receptiveness and initiation of breeding activity. The adults were collected between March 25, 2000 and April 22, 2000, when surface water temperatures in the PSA were 8 to 10 °C. These temperatures represent the ideal environmental "triggers" for the frogs to emerge in the early spring and gather in breeding areas (Gilbert et al. 1994; Hine et al. 1981). Although EPA considered the possibility that the females collected from the target pools were healthy but just experiencing delayed development of their reproductive cells, given the specimen collection time and temperatures, and adult male chorusing, the females would be expected to contain mature oocytes.

  – GE's spring 2003 leopard frog egg mass survey failed to locate significant numbers of egg masses. Of 44 ponds surveyed which were considered to provide suitable leopard frog habitat, only 17 ponds contained any egg masses, with a total of 216 egg masses over the entire PSA. Given that the study area is 10 linear miles of riverine floodplain, this number of egg masses would not be sufficient to ensure a typical level of juvenile recruitment. GE found 2.4 egg masses per hectare (ha) overall and 4.3 egg masses per ha if only ponds with eggs were considered. These numbers are well below the 277 egg masses per ha reported by Hine et al. (1981) or the 58 egg masses per ha reported by Gilbert et al. (1994).

The revised ERA will include a more detailed description of the timing of leopard frog collection in relation to local temperatures and the methods used for collection. Additional life history information on leopard frog breeding activities will be provided. As requested by one Reviewer, the terminology will also be clarified for frogs obtained from Carolina Biological Supply (CBS) and those captured in the uncontaminated areas in the field to control versus reference.

12.B. Wood Frog Population Modeling

The revised ERA will include additional description of the wood frog population modeling, particularly with reference to the parameterization of the model, and further discussion regarding how density-dependence is incorporated into the model. A sensitivity analysis will be used to explore how changing the assumptions regarding density-dependence influences model output. EPA evaluated the GE wood frog study in detail when preparing the ERA, and incorporated the findings into the ERA as appropriate; however, the results of the study produced limited usable data for the population model. In the revised ERA, the use of the results of the GE study will be made more transparent.

The issue of extrapolation of model results to an actual population is often a challenge with population modeling. With amphibians, there is little information in the literature that confirms things such as “normal” ranges for metamorph recruitment, or influence of immigration on the population. These issues, particularly with reference to the potential inconsistency between the population model results and observed wood frogs in the PSA, will be discussed in greater detail in the revised ERA.

13. ASSESSMENT ENDPOINT - FISH

SUMMARY OF ISSUE:

Four Reviewers indicated that a more complete summary of the USGS Phase I and Phase II toxicity studies should be provided in the body of the ERA. Suggestions included:

- A discussion of how observed differences in reproductive endpoints relate to tPCB concentrations measured in adult largemouth bass (i.e., dose-response).
- Acknowledgement of endpoints that were not indicative of effects, to put the observed effects in context, including cases where effects appear to be ameliorated with age.

- More complete presentation of Phase I largemouth bass endocrine and histological endpoints, including reduced steroid hormones and gonadal abnormalities, which are "reliable and appropriate measurement endpoints for reproduction."

Three Reviewers commented on the factor of 4 used to extrapolate from PSA fish to coldwater species below Woods Pond. Although no Reviewer commented that the number was inappropriate, there were questions that the derivation of the value was not sufficiently transparent. One Reviewer commented that an extrapolation factor is synonymous with an uncertainty or application factor.

Two Reviewers commented on the treatment of fish lipid contents, and the assumption that lipids are a controlling factor for PCB accumulation (and toxicity). Two Reviewers indicated the need to explicitly discuss how PCB and TEQ concentrations vary seasonally in relation to lipid content of fish and how this may affect temporal exposure and potentially risk. One Reviewer discussed the uncertainty associated with several fish tissue samples that have low lipid contents (< 0.3%), as described in the RFI (BBL and QEA 2003).

Four Reviewers indicated that additional presentation of fish tissue chemistry data would be helpful. The requests fell into two main categories:

- Mapping of fish collection sites and linkage of exposure data to sampling locations (four Reviewers).

- Additional detail for PCB tissue concentration data (one Reviewer).

Five Reviewers commented on the treatment of morphological deformities in adult fish. Although there was disagreement among Reviewers on the implications of these effects for population health, there were several suggestions to include more information on deformities and disease in PSA fish.

Six Reviewers commented on the need to more fully describe the potential impacts to populations based on field study findings. Specifically, several Reviewers recommended that more analyses and discussion of the observation of skewed population demographics (greater number of old individuals) should be provided.

Three Reviewers commented on the use of the field studies in the overall characterization of risk for this endpoint. Two Reviewers commented that the fish field studies demonstrate a lack of large-scale population effect, rather than an undetermined effect. One Reviewer commented that the field studies do not provide evidence of a healthy population.

RESPONSE:

13.A. USGS Fish Toxicity Studies

Although there is a thorough discussion of the fish toxicity studies and endpoints provided in the USGS reports (attachments to the ERA), EPA concurs with the Reviewers' comments and will provide a more detailed evaluation of these endpoints directly in the revised ERA. The number and type of endpoints exhibiting effects will be compared against the number and type of endpoints that do not show effects, as part of a lines-of-evidence approach.

There was disagreement among Reviewers regarding the ecological significance of some individual-level biological responses (particularly in terms of population-level impacts). The revised ERA will include more discussion of biomarkers such as endocrine and histological endpoints.

13.B. Extrapolation Factor for Coldwater Fish

EPA agrees that a more detailed rationale for the derivation of the extrapolation factor would improve the discussion in the ERA. EPA will amend the revised ERA to provide information on the following:

- Differences in dioxin-like toxicity between warmwater species and trout species, as documented both from the literature and from the site-specific tests.
- Discussion of intraspecies variations in rainbow trout toxicity (i.e., differences among strains).
- Discussion of uncertainty and requirement (if any) for application of uncertainty factors.

The factor of 4 (extrapolation factor) was not an uncertainty factor. When EPA uses the term "uncertainty factor" in the ERA, this term refers to a numeric factor applied to a threshold concentration to ensure that risks are not underestimated, in acknowledgement of uncertainties in the knowledge base. The extrapolation factor, on the other hand, represents a conversion factor based on known information regarding interspecies and inter-strain differences in trout sensitivity. By this definition, an uncertainty factor focuses on what is not known, whereas an extrapolation factor focuses on what is known, about sensitivity differences. This issue is really one of semantics and will be addressed in the revised ERA by more clearly describing the process used in deriving the MATC for coldwater species.

13.C. Lipids in Fish

Linkage of PCB Bioaccumulation to Lipid Content

There is site-specific information on the relationship between lipid content and bioaccumulation of PCBs in PSA fish:

- Analysis of the fish tissue database indicated that lipid content is one of several important controlling factors for PCB bioaccumulation (in addition to fish species and age).
- Bioaccumulation model results (WESTON 2004) and field data (WESTON 2003) suggest that differences in PCB bioaccumulation among species at the same trophic level (e.g., white suckers, brown bullhead, and goldfish) can be explained on the basis of differences in whole body lipid contents.
- An independent evaluation by General Electric of adult largemouth bass PCB concentrations and lipid contents (BBL and QEA 2003) confirmed the findings of the EPA fish tissue collections. This study provides evidence that the relationship between PCB bioaccumulation and lipid content is relatively stable across sampling years.

EPA will summarize this information to provide more discussion regarding lipid normalization.

Seasonal Effects

EPA agrees that seasonal changes in lipid content may in some cases mediate PCB bioaccumulation and toxicity. However, the site-specific field data do not allow for a quantitative assessment of seasonal variations. The vast majority of tissue samples were collected in the late summer and early fall, specifically to avoid the effects of seasonal changes in lipid content and other factors. The USGS Phase I fish toxicity studies implicitly accounted for site-specific changes in lipid composition, since the eggs used in the study were obtained from field-collected fish. EPA will explore whether other data are available to examine the issue of seasonality.

Low Lipid Values

EPA agrees that discussion of the uncertainty associated with low lipid concentrations should be included in the ERA. However, EPA does not believe that the effect of this uncertainty on the characterization of risk for fish is large. The majority of samples with low lipid content identified by GE were fillet samples. The whole body fish tissue PCB concentrations depend mainly upon the concentrations in the offal portion of the fish. This is due to the greater mass of offal samples relative to fillet (approximately 4 four times greater, on average). The independent evaluation of largemouth bass lipid content in 2002 (BBL and QEA 2003) confirmed the results of the EPA sampling.

13.D. Additional Presentation of Fish Tissue Chemistry and Spatial Analysis

EPA will provide a graphical presentation of sampling sites and a summary of tissue concentration data for all individual fish tissue samples.

Because of the large number of PCB congeners reported by the analytical laboratory, EPA believes that a detailed presentation of each congener concentration in every fish tissue sample would not significantly improve the understanding of risk. The congener evaluation (Attachment C.7) investigated whether the congener composition changed significantly across reaches or trophic levels, and the conclusion was that the changes in the congener profile are small. Therefore, presentation of the full congener data set would lengthen an already lengthy document, without commensurate gain in its utility for making risk management decisions. Reconstituted PCB concentrations for congeners have been prepared as part of the bioaccumulation modeling effort, and these concentrations will be provided along with the ERA (in spreadsheet format) for those interested in individual data points. The compact disk provided in ERA Appendix L included congener data for fish tissue.

13.E. Diseased/Deformed Fish and DELTs

EPA agrees that presentation of all information collected on fish health is appropriate, and will make the following changes in the revised ERA:

- Summarize the results of gross evaluations made during the EPA fish tissue collection effort, including incidence of glob-eye and other deformities, erosions, lesions and tumors (DELTs).
- Discuss histological evaluations of growths observed on goldfish collected from the PSA. This evaluation will be based on a report from the USFWS summarizing results of viral and histological assays of Woods Pond fish tissues.

13.F. Population Effects/Population Modeling for Fish

EPA concurs with the comment regarding further exploration of the population-level effects in fish, and will provide the following modifications in the revised ERA:

- A more detailed discussion of the population demographics.
- A discussion of factors, both natural and anthropogenic, that could explain the observed age structure.

13.G. Characterization of Risk

EPA agrees, as noted by one Reviewer, that the presence of self-sustaining fish populations does not equate to a determination of population health. Conversely, the presence of largemouth bass and other species provides evidence that effects such as larval deformities have not translated to readily observable impacts using the field study methods.

14. ASSESSMENT ENDPOINT - INSECTIVOROUS BIRDS

SUMMARY OF ISSUE:

Three Reviewers noted that the tree swallows have a low sensitivity to PCBs compared to most other birds. As a result, Reviewers questioned using this receptor to extrapolate risks to other avian insectivores.

There were divergent opinions among Reviewers regarding how the statistically significant negative relationship between tPCB concentrations in eggs and hatching success of tree swallows be used in the weight-of-evidence assessment for insectivorous birds. Some Reviewers were satisfied with the interpretation in the ERA, i.e., that the relationship was weak and thus the field study "supports the conclusion that tree swallows are not being adversely affected due to exposure to PCBs and TEQ in the PSA." Other Reviewers were not in agreement with this interpretation and had different opinions as to what constituted the correct interpretation. For example, one Reviewer contended that there was no relationship between tPCB concentration in eggs and hatching success. Conversely, another Reviewer suggested that a more appropriate interpretation would be that "swallows are affected but not severely so." Another Reviewer went even further and suggested that the evidence from the field study indicated that "a ranking of 'low risk' for tree swallows is questionable."

Three Reviewers suggested that the tree swallow and American robin have quite different feeding strategies (i.e., tree swallows are aerial insectivores, robins are ground insectivores) and should have had a greater degree of separation than is currently the case in the ERA. Suggestions for degree of separation ranged from:

- Retaining both species in the current chapter and appendix on insectivorous birds, but treating both species separately through the final risk characterization, or
- Treating each species (and species having similar feeding strategies) as different Assessment Endpoints.

RESPONSE:

14.A. Use of Insensitive Bird Species

The tree swallow was selected as a representative species 6 years ago, before much of the current data were available which suggest that this species is not particularly sensitive to PCBs. At this time, there is little information available regarding the sensitivity of potential alternative avian insectivores to PCBs (see Review of Effects literature in ERA Appendix G). Without this information, it is difficult to determine whether other avian insectivores would be at lower or higher risk than tree swallows. Therefore, in the revised ERA, EPA will:

- Acknowledge that tree swallows are relatively insensitive to PCBs compared to other bird species.
- In the conclusions section for insectivorous bird species, characterize risks to the representative species (tree swallow, American robin) and acknowledge the uncertainties in extrapolating these risks to other insectivorous bird species.
- EPA will add a waterfowl (wood duck) receptor to the revised ERA. The waterfowl tissue data will be used to estimate exposure in adults. Contaminant concentrations in wood duck eggs (via maternal transfer) are being analyzed in a limited study in late spring 2004. These data will be incorporated into the revised ERA.

14.B. Hatching Success of Tree Swallows

Section 7.5.2 of the ERA states that,

> "Total PCBs, dioxins, and furans were negatively correlated with hatching success in 1998 and 1999, but the correlations were weak. Hatching success was not correlated with these COCs in 2000, probably because concentrations were reduced in 2000 and because cold weather contributed to poor hatching at all locations."

Virtually identical statements were made on page 33 of the report by Custer (2002). In several other places in the ERA, however, the 2000 results were not mentioned. EPA will correct this omission in the revised ERA.

The ERA concluded that the tree swallow study indicated no evidence of harm from exposure to tPCBs (e.g., Table 7.5-4) or TEQ (e.g., Table 7.5-5). EPA, however, concurs with the Reviewers who suggested that the field study results noted above do indicate some evidence of harm and will modify the revised ERA to reflect this interpretation. EPA continues to believe that the results of the field study indicate low magnitude of effects of tPCBs and TEQ to tree swallows. This interpretation will thus remain unchanged in the revised ERA.

14.C. Tree Swallow and American Robin

EPA concurs with the Reviewers that the species should have a greater degree of separation in the revised ERA. Thus, in the revised ERA, both species will be retained in the chapter and appendix on insectivorous birds, but will be treated separately from the initial conceptual model phase through to the final risk characterization phase.

15. ASSESSMENT ENDPOINT – PISCIVOROUS MAMMALS

SUMMARY OF ISSUE:

Three Reviewers suggested that mink and river otter should have a greater degree of separation than is currently the case in the ERA. One Reviewer suggested that they be separated in the weight of evidence tables "because different kinds and amounts of information were available for each species." Suggestions for degree of separation ranged from:

- Retaining both species in the current chapter and appendix on piscivorous mammals but treating both species separately through the final risk characterization (including weight-of-evidence tables), to

- Treating each species as different Assessment Endpoints.

Two Reviewers requested that EPA address GE's criticism that there is no evidence of a dose-response relationship for survival of kits in the mink feeding study.

Three Reviewers noted that mink are unlikely to feed on goldfish and carp in the PSA, the two species of fish collected for the mink feeding study conducted at Michigan State University. The comment was that goldfish and carp could have different contaminant mixtures and higher PCB concentrations than other fish species consumed by mink in the wild.

Five Reviewers commented on how the jaw lesion and EROD data collected from the mink feeding study were used to estimate risk to mink. Opinions on this issue varied. Some Reviewers thought that the sublethal endpoints such as the jaw lesion endpoint should have been used to a greater extent than was done in the current ERA. Others thought that the current ERA "overstated the significance of the jaw lesions found in kits." Reviewers also noted that the suggestion in the current ERA that the jaw lesions could lead to starvation was speculative.

RESPONSE:

15.A. Representative Piscivorous Mammal Species

EPA concurs with the Reviewers that the species should have a greater degree of separation in the revised ERA. Thus, in the revised ERA, both species will be retained in the chapter and appendix on piscivorous mammals, but will be treated separately from the initial conceptual model phase through to the final risk characterization phase, including the weight-of-evidence tables.

15.B. Mink Feeding Study

Dose-Response in Mink Feeding Study

As described in the report by Bursian et al. (2003) on the mink feeding study, kit survivability was analyzed by analysis of variance (ANOVA) involving the factors treatment and sex (when applicable), with repeated measurements on mink, over another factor, date. SAS PROC MIXED was used to model a first-order autoregressive correlation structure for repeated measurements over dates within mink, as residuals involving measurements taken at adjacent time periods are more likely to be highly correlated than measurements taken further apart in time (Gill 1990). Two-way interactions between treatment, sex, and dates were modeled. After performing the PROC MIXED analysis on kit survivability at birth, three weeks of age and six weeks of age, the decision was made to analyze each time period separately because the data showed possible treatment differences at each time period. It was determined that pooling the least square means for kit survivability across the three time periods was confounding the pair-wise treatment comparisons. Thus, the SAS PROC MIXED command for kit survivability assumed a treatment by date interaction to allow for pair-wise treatment comparison at each individual time period. Kit survivability was percentage (p) data subjected to arcsine, square root transformation $[x=\sin^{-1}(p)]$ prior to statistical analysis. To control for experimental Type 1 error rates, a Tukey's honestly significant difference procedure was used to test comparisons between means based on the total number of pairwise comparisons. EPA believes that the above described statistical analysis is a reasonable methodology given the design of the study. The analysis indicated that the highest PCB dose treatment had a significant effect on kit survivability at six weeks of age.

EPA will review the analysis conducted by Bursian et al. (2003) in the revised ERA. In addition, EPA will conduct a regression analysis and also consider other statistical techniques. The regression analysis will allow for a better discussion of the relationship between the dose-response from the mink feeding study and the dose-response presented in the ERA from published Aroclor 1254 feeding studies, as these analyses would then be more directly comparable. This will provide for better use of the mink feeding study data in the risk characterization, as was noted by one Reviewer.

Fish Species in Mink Diet

To address this issue, one Reviewer recommended including an evaluation of how the contaminant mixture in the fish used for the diet compares to that for the fish assumed to represent the diet of mink in the PSA. EPA already conducted this evaluation in the ERA (presented in Appendix I, Section I.3.2.6). This evaluation concluded that "the PCB composition in fish from the feeding study can be treated as similar to that in fish used in the exposure analyses and the results from both studies are directly comparable" (page I-51).

EPA also notes that the TEQ analyses account for any minor differences in contaminant mixture, if they existed, between the goldfish and carp collected for the feeding study and fish consumed by mink in the PSA. The TEQ calculation is a weighted (by toxicity relative to 2,3,7,8-TCDD) sum of concentrations of congeners that exhibit dioxin-like toxicity. Total PCBs were also measured in each of the dietary treatments, and this measure thus accounts for any differences in tPCB concentrations that might exist between goldfish and carp and other fish species in the PSA.

Sub-Lethal Endpoints in Mink Feeding Study

The mink jaw lesions have been the subject of continuing study since the release of the ERA. In a recent presentation at Society of Environmental Toxicology and Chemistry (SETAC) (Beckett et al. 2003), it was noted that the data at this point appear to indicate that the lesions are in fact neoplastic. The pathologist who analyzed the samples from the mink feeding study conducted by Michigan State University, Dr. Ben Yamini, has also indicated that the lesions appear to be cancerous. EPA continues to believe that the jaw lesion and EROD data provide information that can be used to help characterize risks to mink. In the revised ERA, EPA will:

- Consider the jaw lesion and EROD data in developing the risk characterization for mink.
- Discuss the significance of recent studies that indicate the observed jaw lesions in mink may be neoplastic, and the potential for such lesions to negatively impact feeding.
- Discuss the significance of recent studies that link the jaw lesions observed in the feeding study with similar lesions that have been observed in wild mink in the Kalamazoo River basin.

16. ASSESSMENT ENDPOINT – OMNIVOROUS/CARNIVOROUS MAMMALS

SUMMARY OF ISSUE:

Six Reviewers suggested that short-tailed shrew and red fox should have had a greater degree of separation than is currently the case in the ERA. In general, the Reviewers felt that more separation was required because the same data were not available for each receptor and/or because the shrew and fox represent different exposure pathways and life histories. Suggestions for degree of separation ranged from:

- Retaining both species in the current chapter and appendix on omnivorous and carnivorous mammals but treating both species separately through the final risk characterization (including weight of evidence tables), to
- Treating each species as different Assessment Endpoints.

All Reviewers commented on the discrepancy between EPA's analysis of the shrew field study and the subsequent re-analysis performed by Dr. Boonstra on behalf of GE. The EPA analysis indicated a significant relationship between the concentration of tPCBs in soil and survival of male and female shrews from summer to fall in the PSA. The GE analysis did not find a significant relationship despite "using the methods that the EPA used (probit analysis) and their values for soil concentrations" (Boonstra 2003, Attachment R from "Comments of General Electric Company on the Ecological Risk Assessment for the General Electric/Housatonic River Site, Rest of River [July 2003 Draft]"). This discrepancy produced a range of responses from Reviewers including:

- "Given the dependence of the statistical significance on subtle differences between two (seemingly) appropriate statistical methods, the most robust conclusion that can be made from this study is that the response is borderline."
- "The GE analysis was not appropriate for this particular data set."
- [The EPA analysis] "indicates a high level of potential harm."
- "The results for the EPA and GE data sets should be compared and aligned."
- "The Boonstra analyses should be included in the ERA and the differences between the two approaches and their results should be discussed."
- "EPA and GE settle on one mutually agreeable approach and use it in the ERA."
- "Both methods should probably be reported and discussed in the final ERA."

RESPONSE:

16.A. Representative Omnivorous and Carnivorous Mammal Species

EPA concurs with the Reviewers that the species should have a greater degree of separation in the revised ERA. Thus, in the revised ERA, both species will be retained in the chapter and appendix on omnivorous and carnivorous mammals, but will be treated separately from the initial conceptual model phase through to the final risk characterization phase, including the weight-of-evidence tables.

16.B. Statistical Analysis of GE Shrew Study

Discussions during the Peer Review between GE and its consultants (primarily Dr. Rudy Boonstra) and EPA and its consultants (primarily Dr. Dwayne Moore and Dr. Scott Ferson) indicated that the analyses performed by Dr. Boonstra and those conducted by EPA used different statistical methodologies rather than as was stated by GE in their presentation that

> "Dr. Boonstra reanalyzed data using EPA's soil PCB concentrations and EPA's statistical technique, but found no significant relationship with survival. (Details in Boonstra comments [Attach. R to GE comments]). EPA erroneously concluded that this study showed "undetermined" evidence of harm. In fact, it showed <u>no</u> evidence of harm."

Rather than use the same methods, the analysis by Dr. Boonstra did not weight each of the grids according to sample size, as was done in the EPA analysis. Dr. Boonstra also excluded one grid in the analysis of the survival of male shrews because the grid had a sample size of one. EPA included this grid; however, it had a very low weighting in the analysis because of the small sample size. Both analyses used a probit transformation to transform survivorship prior to conducting a linear regression analysis.[1] The EPA analysis assumed an underlying binomial error distribution in their analysis because survival is a quantal endpoint. According to discussions held during the Peer Review, Dr. Boonstra used the methods described in the first edition of *Biometry* (Sokal and Rohlf 1969) – this methodology is a least squares analysis which means that the underlying distribution was assumed to be normal (an incorrect assumption for quantal data – Bailer and Oris 1997).

EPA concurs with the comment that the conclusion from this study is that the dose-response relationships were not strong. This is the conclusion that was originally arrived at by EPA in the ERA, i.e., "Although the results of the analyses indicated a significant relationship between soil concentrations of tPCBs and shrew survival, the confidence limits indicate that the relationships are not strong" (page J-57) and "The slope of the regression models is not steep indicating that survival was only slightly reduced at the 'high' contaminated grids compared to the 'low' contaminated grids" (page J-58). EPA does not concur with the comment that the analyses indicate a high level of potential harm for that study location. However, because Reviewers commented that individual site responses should be placed in context with overall site concentrations, the revised ERA will include a presentation of the re-analysis conducted by Dr. Boonstra subsequent to the release of the current draft of the ERA, and an evaluation of the effects in the context of the data from the other areas in the PSA.

---

[1] Two Reviewers noted that the Bailer and Oris (1997) method used by EPA can be used to model any shape of dose-response relationship by using a quadratic model equation instead of a linear model equation following probit transformation of the data. Although true, EPA used a linear model, not a quadratic model, in its analysis of the shrew field study.

OVERVIEW COMMENTS

Valery Forbes:

- The assessment endpoints should be redefined so that they are more consistent with general EPA practice and so that they more accurately reflect the protection goals that were actually used in this ecological risk assessment (i.e., long-term persistence of local receptor populations).

RESPONSE O-VF-1

Please refer to the response to General Issue 2.

- More transparency and consistency is needed in describing the WOE approach. Describing the process, or parts of it, using the phrase 'best professional judgment' should be avoided. More care should be taken in combining lines of evidence that are not independent. The WOE summary tables should be modified so that they are more self-explanatory and less ambiguous.

RESPONSE O-VF-2

Please refer to the response to General Issue 8.A.

- More detailed and consistent descriptions of the statistical methods used should be provided in those parts of the ERA where data are presented (the reader should not be referred to the original article to find out what kind of statistical test was used). Both statistical significance and effect size should be reported and considered in the risk characterization.

RESPONSE O-VF-3

Please refer to the response to General Issue 1.

- Interpretation of HQ results needs to be refined. Both the magnitude of the maximum HQ as well as a measure of the probability (or proportion of samples) exceeding an HQ of 1 (or 10, or 100 as appropriate) should be included; it should be clear whether the spread in the HQs derives from variability in exposure (the numerator), variability in effects (the denominator), or both. Given that HQs provide a rather coarse measure of risk, differences in HQs of less than an order of magnitude should not be considered as indicating differences in risk.

RESPONSE O-VF-4

Please refer to the responses to General Issues 7 (interpretation of HQs) and 1.C (probabilistic HQs). EPA believes that when HQs are calculated based on screening criteria, such as sediment quality values, HQs may provide a conservative measure of risk, but with associated uncertainty. However, when HQs are calculated based on site-specific thresholds (as in Section 12), differences of less than an order of magnitude can indicate differences in risk magnitude, although the difference is not necessarily proportional to the numeric expression.

- The ERA should avoid use of value-laden terms to describe risk (e.g., catastrophic, unacceptable), and instead aim to quantify the likelihood and degree of impact in objective terms as best as possible.

RESPONSE O-VF-5

Please refer to the response to General Issue 7.

- The panel identified a number of studies/analyses that could have been done in the context of the risk assessment. I do not recommend that completion of the ERA be delayed in order to include more studies in it. However, given that an important output of the ERA is the identification and quantification of important sources of uncertainty, I would strongly recommend that actions taken on the basis of the ERA include both consideration of remediation alternatives as well as additional, highly focused, studies/ analyses designed to address the most important uncertainties identified in the ERA.

RESPONSE O-VF-6

Focused sampling may be undertaken following completion of the ERA to support the design of the remediation plan (e.g., to reduce sources of uncertainty, identify hot spots) or to determine whether remediation actions are achieving their goals. Long-term monitoring will be a component of any action.

- Serious consideration should be given to restructuring the ERA to limit the redundancy between the Assessment Endpoint Chapters in the main document and the relevant Appendices in which all of the details are found. In my view the Endpoint Chapters provide too much information for the casual reader and not enough for the interested expert. These could be deleted from the main document since all of the information they contain is provided in the Appendices. A series of maps that overlay sampling sites for exposure estimates and sampling sites for the various effects estimates would be a very helpful addition to the document.

RESPONSE O-VF-7

Please refer to the response to General Issue 1.

- According to EPA guidance, ERAs should use site specific studies wherever possible. Unfortunately many of the field studies performed in the context of the present ERA suffered from weaknesses related to one or more of the following: no reference sites; small sample sizes; short study durations (e.g., one reproductive season); they addressed a question that did not lend itself easily to incorporation in the WOE (e.g., is species X reproducing in the PSA, yes or no?). This is extremely unfortunate since the potential strength of site specific field studies is that they deal directly with mixtures of chemicals (and other stressors) present at the study site and should therefore have less uncertainty (and weigh more heavily) than laboratory studies or models. I would recommend that EPA and GE work together toward developing some guidance on the appropriate design of field studies for use in these kinds of ERAs in the interest of improving future projects of this nature.

RESPONSE O-VF-8

EPA will consider mechanisms to communicate what has been learned at this site to other parties in the future.

- It would be extremely valuable if the EPA and GE could jointly compile a document that highlights the lessons learned from the Housatonic risk assessment project in a format that could provide guidance for the successful conduct of future risk assessments of this kind.

RESPONSE O-VF-9

See Response O-VF-8, above.

James T. Oris:

The ecological characterization and the ecological risk assessment presented for the Housatonic River ecosystem is a very impressive set of documents. At the surface, they are extremely comprehensive and thoughtfully compiled. However, there are numerous aspects of the organization of the documents that should be considered prior to finalization of the ERA. Two areas of revision would be considered major revisions to undertake.

Comment B1. The first area is of document organization. Currently the document is extremely large and contains an enormous amount of information. It would serve the Agency, GE, and the public if a significant effort were exerted to organize the document better. The document is organized with a general executive summary, specific chapters with more details, and detailed appendices for each chapter, and then detailed appendices that cover methods and analysis that overlap among chapters/appendices. Based on my reading, the specific chapters and the appendices have a tremendous amount of overlap and redundancy. This was explained as being necessary because the document was meant to serve multiple purposes and audiences. I submit that most laypersons will only read a 1 - 10 page brief summary (e.g., as in the glossy documents meant for the public). Many highly interested persons with a variety of educational backgrounds will be able to work through and understand the executive summary. In my opinion, only technically competent people will read the specific chapters or be interested in the appendices. Therefore, I see no reason to have both chapters and their associated detailed appendices. I submit that the document would be best organized by having all of the detail within the specific chapters. Those persons reading the specific chapters will be, because of their professional training, accustomed to skimming sections for appropriate detail. They also (like myself) will be frustrated by repeated cross-referencing to different parts of the documents, and then having to search through redundant information to find the one or two paragraphs that have those details. A 'redundancy analysis' or 'compare documents' exercise between specific chapters and their appendices will show that a remarkable reduction in document length and a remarkable increase in organizational clarity could be achieved by putting the details in the chapters and eliminated most of the appendices. This could be achieved with a minimal disturbance to the readability of the document. There are some appendices that should not be eliminated, however. These would include sections that are redundant among chapters (e.g., methodology on setting values for non-detects).

RESPONSE O-JO-1

Please refer to the response to General Issue 1.

Comment B.2. The second area is of document transparency. I found it difficult to connect all the pieces of the ERA together into one coherent overview or conceptual framework. In addition, a significant effort needs to be exerted to eliminate confusing statements (e.g., see comment 1.1 below), differential application of methods (e.g., risk characterization of fish and invertebrates versus birds and wildlife), and differential use of terminology (e.g., risk terminology). The organization of the document (as outlined above) also hinders the transparency of the document (it is hard to find information you are seeking because it is in so many different places). Finally, the least transparent portion of the document is the description and application of the weight of evidence approach. How final weightings were assigned, how each line of evidence was used, why sometimes lines were eliminated, and how final rankings were decided based on all lines and weights combined is not clear in any of the sections. This is the case even after reading the weight of evidence papers in the literature. Perhaps this is just my own confusion, but I think it would be useful to provide a detailed, step-by-step example of how the method was utilized for one of the receptors and then ensure that the method is applied consistently throughout the document. The weight of evidence approach, with the current level of explanation, is open (in my opinion) to criticism of subjectivity in the conduct of the risk assessment. I do not believe that there was bias in judgments by EPA, but this is an area that can be considered a weak point in the document.

RESPONSE O-JO-2

Please refer to the responses to General Issues 1 and 8.A.

An example of the WOE unclarity can be found using an analysis of the assignments of weights in Table D.4-1 for the invertebrate assessment. Below is a reproduction of that table, objectively assigning point values to each weighting classification, using three different averaging methods to calculate weighting classification scores, and then a comparison to the EPA assigned overall weightings in the original table. While there is detailed explanation for why each item in the table received a particular weighting, the lack of transparency in the EPA approach is that there is insufficient explanation of how the final, overall endpoint weighting value is assigned. In my analysis below, the first averaging method assumes each category of weighting values has equal weight, calculates an average within each category and then an overall average of the three categories. The second averaging method assumes that each of the 10 subcategories has equal weight and the average is calculated over all of these values (i.e., sum of raw weightings over a column divided by 10). The third method calculates average weighting values within a category, re-ranks the weight, and then averages the value of the new ranking. Either method 1 or method 3 are probably the most appropriate because they provide even estimates of weights across the three categories since there are different numbers of subcategories within the three (i.e., Relationship between Assessment and Measurement endpoints (3 subcategories), Data Quality (1 subcategory), Study Design (6 subcategories). Regardless of how these are calculated, there is a fair amount of discordance for several of the endpoint group weightings between my analysis and the EPA-assigned values. This is especially true for endpoint groups C1, C2, C3, and T3 (nearly 50% of the endpoint groupings). I am not sure how this will affect the outcome of the risk characterization for invertebrates, and I did not take the time to evaluate each of the weighting tables. It is evident, though, that there needs to be more objectivity, better explanation, and more clarity to this analysis than is currently present in the document.

Analysis of Table D.4-1, Weighting of Measurement Endpoints.

Rankings within single cells: 1=Low, 2=Low/Mod, 3=Mod, 4=Mod/High, 5=High
Rankings of averages: 1-1.66=Low, 1.67-2.66=Low/Mod, 2.67-3.66=Mod, 3.67-4.66=Mod/High, 4.67-5.00=High
Abbreviations: L=Low, LM=Low/Mod, M=Mod, MH=Mod/High, H=High

Category: Endpoint	Groups: C1	C2	C3	T1	T2	T3	B1	B2	B3
Meas vs. Assess Endpoint:									
1	1.00	1.00	3.00	4.00	4.00	3.00	2.00	2.00	2.00
2	1.00	2.00	3.00	4.00	4.00	4.00	2.00	2.00	2.00
3	1.00	3.00	3.00	5.00	5.00	4.00	2.00	2.00	2.00
Avg of Cat.	1.00	2.00	3.00	4.33	4.33	3.67	2.00	2.00	2.00
Rank of Cat.	L	LM	M	MH	MH	MH	LM	LM	LM
Score of Cat.	1.00	2.00	3.00	4.00	4.00	4.00	2.00	2.00	2.00
Data Quality:									
4	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
Avg of Cat.	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
Rank of Cat.	H	H	H	H	H	H	H	H	H
Score of Cat.	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00	5.00
Study Design:									
5	2.00	2.00	2.00	4.00	4.00	4.00	5.00	5.00	5.00
6	2.00	2.00	2.00	5.00	4.00	5.00	2.00	2.00	2.00
7	4.00	5.00	3.00	3.00	4.00	1.00	4.00	4.00	4.00
8	5.00	5.00	4.00	3.00	3.00	3.00	3.00	3.00	3.00
9	4.00	4.00	4.00	5.00	5.00	4.00	5.00	4.00	3.00
10	3.00	5.00	5.00	5.00	4.00	4.00	3.00	5.00	5.00
Avg of Cat.	3.33	3.83	3.33	4.17	4.00	3.50	3.67	3.83	3.67
Rank of Cat.	M	MH	M	MH	MH	M	MH	MH	MH
Score of Cat.	3.00	4.00	3.00	4.00	4.00	3.00	4.00	4.00	4.00
Avg of (Avg of Cat.)	3.11	3.61	3.78	4.50	4.44	4.06	3.56	3.61	3.56
	M	M	MH	MH	MH	MH	M	M	M
Overall Avg-no Cats avg'd	2.80	3.40	3.40	4.30	4.20	3.70	3.30	3.40	3.30
	M	M	M	MH	MH	MH	M	M	M
Avg of Cat. Score	3.00	3.67	3.67	4.33	4.33	4.00	3.67	3.67	3.67
	M	MH	MH	MH	MH	MH	MH	MH	MH
EPA Rating in D.4-1	LM	LM	M	MH	MH	M	M	M	M

RESPONSE O-JO-3

Please refer to the response to General Issue 8.A.

Mary Ann Ottinger:

I appreciate the opportunity to participate in the peer review of the Ecological Risk Assessment (ERA) for the Rest of the Housatonic River conducted by the Environmental Protection Agency (EPA). The experience has been challenging, interesting, and instructive. As a behavioral endocrinologist and an avian biologist, I have learned a tremendous amount about field evaluation of risk from PCBs. Much of my work has been devoted to avian species, and in assessing the consequences of embryonic and life time exposure to endocrine disrupting chemicals (EDCs). These studies have focused on establishing reliable indices of EDC exposure and have addressed questions of species and age-related sensitivity to the chemicals. In addition, the research by colleagues at Patuxent Wildlife Research Center and many other laboratories have provided a growing body of research on the consequences of EDC exposure in birds and on the basic biology of birds. We now understand many of the mechanisms that regulate reproductive endocrine and behavioral responses in both laboratory models and in wild birds, specifically those that are likely to be impacted by EDCs. It is from these vantage points that I have read the information provided and formulated responses to the charge questions, with a focus on the relevance of measurements for assessment of long-term consequences for field populations.

Our charge was to review the EPA's ERA in terms of EPA policy and guidance; protocols used in the risk assessment, interpretation of findings from the studies, and the ERA conclusions. The ERA focuses on the portion of the river from the confluence of the river, 2 miles below the GE facility to Woods Pond dam and associated floodplain, and is termed the Primary Study Area (PSA). A great deal of information was provided to the Panel, which reflects a considerable body of work and a great deal of effort by many professionals over a number of years. My responses are based on these written documents and on the information that was provided to us at the two meetings and tour of the site at Pittsfield, Massachusetts in October and December 2003. Those meetings were extremely helpful in providing information about the site from a chronological aspect as well as to see how the site appears today, including areas undergoing active remediation. I found the second meeting helpful in that the scope of numerous studies were placed in context for the panel. Both the EPA and GE teams are to be complemented for their extensive work on this project and the EPA team in collaboration with the GE team did an excellent job of ordering the material for us in a logical manner.

Several points deserve further attention. This is clearly an extremely complex site, with relevance for the ecosystem including the human impact. As such, the ERA includes multifaceted assessment endpoints and summarizes the risks based on the "weight of evidence". This approach, while appropriate, tends to oversimplify the consequences of the long-term PCB exposure by virtue of repeated data reduction, even though the models take into account the probabilities and uncertainties. Some of my comments are directed at potential error in conclusions due to the need to summarize the findings to the point that salient data become overwhelmed. Although appropriate in the "weight of evidence" approach, my concern is that the risk to selected groups of wildlife is underestimated or minimized.

RESPONSE O-MO-1

Whenever possible, EPA tried to use a conservative, yet reasonable, approach when assessing risk to selected receptors. However, due to the uncertainty associated with methodologies employed to evaluate risk, there are instances where EPA may have underestimated potential risks to target receptors. EPA will review the discussion of uncertainty associated with the avian receptor analyses and will adjust the uncertainty discussion to address the potential to underestimate avian risks.

Detailed comments are provided below in response to the questions provided in the Charge to the Panel.

Brad Sample:

Attached are my peer review comments on the US EPAs series of documents entitled: "Ecological Risk Assessment for the General Electric (GE)/Housatonic River Site, Rest of River". My review has been conducted in accordance with the charge for the ecological risk assessment peer review, with each specific question individually addressed. I scanned all materials provided, but, as my expertise is in wildlife toxicology and risk assessment, I focused my attention (and the bulk of my comments) on these sections. Comments on the sections dealing with benthic invertebrates, fish, and amphibians are more limited. My review focused on the materials presented in or otherwise associated with the appendices. I did little more than a cursory review of the associated chapters in the main text.

Both EPA and GE should be commended for the extensive time and effort they have clearly invested in this assessment. In addition to possessing one of the most extensive ecological risk assessment datasets I have seen, this report, and voluminous associated supplementary material, provide an example of the application of many state-of-the-art methods in ecological risk assessment.

Before addressing the charge questions, I have a couple of general presentation-related comments. Although the report is very well written, given the complex data and highly technical material, and because of the massive volume of information evaluated and presented, it can be very difficult to follow. The presentation style used, in which there was a main text that presented a somewhat abbreviated version of the analyses, with the full analyses presented in an appendix, added unnecessary length to the document. I would recommend that the materials in Chapters 3 through 5 and 7 through 11 be replaced by the more detailed discussions presented in the appendices for each respective assessment endpoint.

RESPONSE O-BS-1

Please refer to the response to General Issue 1.A.

In addition, because the data used in the assessment are presented in various volumes and appendices, it was very difficult to locate specific data used for each assessment endpoint. To help correct this, I would recommend that all data (i.e., abiotic and biotic chemical concentration data used to estimate exposure, including maps displaying where the data were collected in relation to each other; plus results of field and laboratory studies conducted by both EPA and GE) used for the evaluation of a particular assessment endpoint be specifically reported in association with that evaluation. Although this could result in the repeated presentation of some data, I think the increased transparency that would result would offset any additional length.

RESPONSE O-BS-2

Please refer to the response to General Issue 1.B.

Specific responses to the peer-review charge questions are presented below.

Ralph G. Stahl:

I have read the entire ERA (Vols 1&2), all of the Appendices, all of the public comments on the ERA, and skimmed the pre-ERA materials and the RFI information. However, there simply was not time sufficient to review all of the documents in detail, nor would that be necessary to render a scientific opinion opposite the charge to the panelists. I participated, by conference phone, in the December 2003 and January 2004 public meetings. I appreciate the opportunity to provide comments and suggestions on the ERA and hope that these will be useful to GE, EPA and the authors of the ERA in subsequent revisions of the final product.

Before I begin my comments, I would like to suggest changes to this particular peer review process. First, the peer review would be much more effective if the panelists were able to communicate among ourselves during and outside of the formal meetings, etc., and in a deliberative fashion. The ability to learn from the strengths each of us brings to the review would be very beneficial to our abilities to formulate more effective suggestions for improving the ERA.

Second, the format of the meeting with GE, EPA and their contractors on December 18, 2003 did not lend itself to the type of scientific dialog that most peer reviews are noted for. I have served on various national panels over the years but have yet to encounter a format such as was used on December 18. Having to formulate questions in lieu of a didactic discussion seemed to constrain the dialog. A similar format was used for the public meetings in January, and once again I felt that our dialog was constrained.

One additional suggestion. If there are similar reviews specified by the consent agreement, I believe it would be much more effective to have provided the documents first and then schedule a series of conference calls for the panelists to discuss their observations with GE and EPA, before holding any public meetings. Even better, having the documents for at least 3 months before any public meetings would lead to a more thorough reading especially if the materials are similar in volume to the current ERA, Appendices, etc. On the other hand, I am cognizant of the need to get the review done in a set period of time, and recognize that much of the peer review for the site activities has been completed.

RESPONSE O-RS-1

The process for the Peer Reviews for the Housatonic River was negotiated as part of the Consent Decree between GE, EPA, and other federal, state and local agencies and groups. The amount of dialogue between the Reviewers and EPA and GE was deliberately limited to specific exchanges on questions of fact for the purposes of clarification, other than the one opportunity for EPA to present the materials, and the opportunities for GE and other members of the public to provide both written and oral comment on the documents to the Reviewers. This was to ensure that the exchange between the Reviewers and all parties was fair.

The Peer Review is a 13-week process from receiving the documents to the Public Meeting. The documents were distributed in week 1; in week 2, there was an Introductory Session for the Reviewers at the site with the Managing Contractor, Neutral, EPA, and GE; in week 4, the Reviewers had the opportunity to submit questions of any nature on the document; in week 10 (the November meeting), EPA provided a presentation of the document and the Reviewers were not restricted in the nature of the questions that could be asked and the Reviewers could discuss their observations with EPA and GE at that meeting.

The Reviewers were not allowed to deliberate "offline," so that all interested parties could hear all of the discussions, which were reserved for the public meeting.

It was unfortunate that this Reviewer had other commitments that placed limitations on participation in some components of the review. However, the process, as described in Appendix J of the Consent Decree, in fact is comparable to what was described by the Reviewer.

Overall, Summary Comments

Overall, the ecological risk assessment (ERA) for the Housatonic River, rest of river, is a substantial amount of work and it is evident that GE, EPA and the authors of the document have invested a significant amount of time and effort to complete this draft. Such a large volume of information was a challenge to review effectively in the time provided.

In general the text was well written and sufficiently detailed although the frequent reference to appendices for more detailed discussion of specific topics was troubling given that the ERA (Vols 1 & 2) was already over 900 pages in length. When faced with such a large document already, the reader is not particularly enthusiastic about being tasked with reading yet more text in an appendix in yet another document. This was a bit tedious but did not detract substantially from the ERA. Given the volume of information generated for the ERA, I do not have a suggestion on how to alleviate this problem.

RESPONSE O-RS-2

Please refer to the response to General Issue 1.

The initial screening of COCs for this ERA is one area where additional discussion among the ERA authors would be useful. After reading parts of the pre-ERA and the ERA itself, it is not fully obvious that tPCBs and TEQ caused all of the toxicity or developmental abnormalities observed in the benthic community, fish and the frogs. Inorganic mercury for example is also capable of causing skeletal and other malformations in finfish, yet mercury was given little discussion within the ERA. There is some evidence that methylmercury may also cause endocrine modulation in some fish (Drevnick & Sandheinrich 2003) In addition, the statement that PAHs were not considered for their potential effects on fish because they are readily metabolized seems to be a major overstatement. I have not read the source paper cited for this statement, but it is difficult to agree fully that benzo(g,h,i)perylene and other high molecular weight PAHs are readily metabolized by fish.

RESPONSE O-RS-3

Please refer to the response to General Issue 3.

In my opinion, it would be much better to have indicated that given the situation on the Housatonic River, tPCBs and TEQ would be the COCs simply on the basis of practicality. The screening process for determining which COCs would be carried into the ERA was effective for the most part, although there are some concerns with the reasoning expressed in the text. I agree that pesticides were not an issue at this site given the analytical chemistry results, and the apparent absence of a pesticide manufacturing activity in the area. I am less convinced however, that mercury and perhaps PAHs were not involved in some or part of the observed toxicity and developmental abnormalities, yet these were not given much credence in the document. Whether or not they originated from site activities is not evident from the ERA or supportive materials. Nevertheless, although it is not mentioned in the ERA (nor should it be) it is obvious that a final remedy that addresses the tPCBs and the TEQ is likely to also address potential risks posed by PAHs and other contaminants in the system.

RESPONSE O-RS-4

Please refer to the response to General Issue 3.

There were some figures and some tables which were difficult to read and/or understand. This is particularly the case where the axes and scales, and labels for the specific plots, were not clearly legible. I have noted these in my more detailed comments. Other panel members have also indicated the need for more detailed maps which illustrate the synoptic sampling of sediments, floodplain soils, surface water and biological tissues. I support this comment and suggest that EPA consider developing maps which more clearly illustrate the combined chemical, toxicological and biological results.

RESPONSE O-RS-5

Please refer to the response to General Issue 1.

Some relatively new approaches to data analysis were provided in this ERA and the individual scientists who conducted these analysis on behalf of EPA are recognized for their expertise in this area. For those schooled in a less robust form of data analysis, it was pleasing and challenging to read through the discussion of using propagation tools and bounding analyses to more fully describe the data as well as the uncertainties within those data sets. Despite the assumptions that underlay the analyses, I think this is a strength of the current ERA. This work is, without question, breaking new ground in ERA and in so doing will no doubt push the science forward. Even so, there were some sections in the ERA where the discussion was much too detailed and the various manipulations of data so complicated that it was not easy to follow the logic. I have noted these in my detailed comments.

The use of the Massachusetts Weight-of-Evidence process so broadly in this ERA is not something I have observed previously in an ERA context. When Dr. Menzie worked with the group to develop the process, I'm not sure any of the participants appreciated that it could be used in this manner. I found that it was difficult to understand just how much professional judgment went into the final assignments in the matrix, and how much "weight" was given to professional judgment compared to the other endpoints in the matrix. After attempting to discern how many empty, partial or full circles it took to achieve a certain final assignment, only then did it become clear that professional judgment, more or less, was the final arbiter of the assignment. Fundamentally it appeared that toxicological data tended to override field data, but not in all cases. The chemical data did not appear to be given a great deal of weight except in the case of benthic invertebrates, where the chemical data were a line of evidence that showed "impacts" throughout the PSA. In most cases the toxicological and biological field data were given greater weight even when there were problems in study design, execution or results (depredation of nests for example). As GE, EPA and the authors of this ERA fully appreciate, it is seldom the case in environmental studies that all lines of evidence are fully concordant. This leads to the use of professional judgment, and its use was in evidence throughout this ERA.

RESPONSE O-RS-6

Please refer to the response to General Issue 8.

I strongly recommend that the authors of the ERA be more forthcoming and descriptive with their use of professional judgment in the WOE discussion so that the reader is not confused about how the final assignments of risk were determined. In this regard, I found the evaluations at the end of each of the assessment endpoints to be inconsistent with one another, and not wholly objective. Others on the peer review panel have noted this in their verbal and written comments. If anything, a high degree of subjectivity is apparent in the final assignment of risk. Whether this can be remedied by clarifying more fully the use of professional judgment remains to be seen.

RESPONSE O-RS-7

Please refer to the response to General Issue 8.

As one of several peer reviewers of the EPA's Ecological Risk Assessment Guidelines in 1998, and the ERA guidelines for Superfund in 1997, I do not recall any codification of the Massachusetts WOE approach. A general discussion of the "weight-of-evidence" is part of the guidelines, but there is no specific recommendation for a particular methodology. Thus its (WOE) widespread application within this ERA would appear to be counter to, or at least outside of existing EPA guidelines. Otherwise, the ERA in general appeared to follow the intent, if not the specifics, of the two EPA guidelines noted above.

RESPONSE O-RS-8

Please refer to the response to General Issue 8. While the Massachusetts WOE approach (Menzie et al., 1996) is not specifically mentioned in EPA guidance, it is an accepted approach that has been used at numerous sites throughout EPA New England Region (Johnston et al. 2002) and fulfills the principles described in the EPA guidance because it uses an established technique.

I would caution however that simply following guidelines does not guarantee a high quality ERA. While it might be useful to have followed guidelines so others understand the basis for how the ERA was conducted, it should not be used as prima facie evidence of providing a high quality end product. The quality and scientific merit come of themselves, not as a result of following guidelines.

Another area where additional discussion might be beneficial centers on the "causative" agents believed to be responsible for the observed toxicity and developmental abnormalities. In that regard, I believe that the authors of the ERA did not consider accurately those substances which might have also elicited some of the same developmental or morphological malformations. I believe the authors would benefit from a review of the EPA's stressor identification guidelines (USEPA 2000). This could be useful in the Problem Formulation discussion to provide support to the statements made later in the ERA concerning tPCBs and TEQ being the primary causative agents for the observed malformations in frogs and fish, and toxicity in some samples. I believe the initial manuscript, by Glen Fox of the Canadian Wildlife Service, for the term "ecoepidemiology" may or may not have been cited. If not, it would be helpful to review this document as well (Fox 1991) and reiterate in the ERA the criteria applicable to assigning causality. See also (Diamond & Serveiss 2001)

RESPONSE O-RS-9

EPA agrees to further review these documents for potential incorporation in the final ERA.

I also reiterate my earlier comment that the ERA is not the document whereby the finding of an unacceptable risk is made. This is a risk management decision, arrived at after reviewing all the pertinent information in addition to the ERA ((Stahl et al. 2001; Pittinger et al. 2001)).

RESPONSE O-RS-10

EPA will remove the terms acceptable or unacceptable risk in the revised ERA.

Summary of Recommendations

1. EPA should consider additional characterization of the sediments located behind dams in Connecticut that are in closest proximity to the PSA.

RESPONSE O-RS-11

Please refer to the response to General Issue 5.

2. In support of the comment made repeatedly during the January 2004 public meetings, it is strongly recommended that the authors of the ERA provide detailed maps that clearly illustrate the synoptic sediment, water column and biological tissue sampling and results.

RESPONSE O-RS-12

Please refer to the response to General Issue 1.

3. The authors of the ERA should more fully explain the reasoning for not including such receptors as the dragonfly and waterfowl.

RESPONSE O-RS-13

Please refer to the response to General Issue 4.

4. The authors of the ERA should temper their statements with respect to causation, tPCBs, TEQ and the observed impacts on aquatic receptors. Other COCs noted in the sediments are known to elicit some of the same developmental malformations attributed to PCBs and TEQ.

RESPONSE O-RS-14

Please refer to the response to General Issue 3.

5. The authors of the ERA should refer to the EPA Stressor Identification Guidance (USEPA 2000) and incorporate the tenets of that guidance into the Problem Formulation section of the ERA.

RESPONSE O-RS-15

Please refer to Response O-RS-9.

6. The authors of the ERA should re-word the assessment endpoints to reflect more specifically those attributes of the assessment endpoint (or receptor) that are most valued and worthy of protection. The current assessment endpoints are overly broad and thereby lead to difficulty in interpretation of and linkage to the measurement endpoint results.

RESPONSE O-RS-16

Please refer to the response to General Issue 2.

7. The authors of the ERA should be more forthcoming and descriptive with their use of professional judgment in the WOE discussion so that the reader is not confused about how the final assignments were determined. It is clear that a large component of the final assignment of risk is directly related to professional judgment. The final assignment of risk was not fully consistent nor objective among the 8 assessment endpoints.

RESPONSE O-RS-17

Please refer to the response to General Issue 8.

8. The term “ecologically significant” is used in various places throughout the ERA; however, this term is not defined adequately within the document. The authors of the ERA should develop and clearly articulate the definition for “ecologically significant”.

RESPONSE O-RS-18

EPA concurs with this comment and will clarify the use of the term “ecologically significant” in the revised ERA. In the current and revised ERA, “ecologically significant” effects are those that are directly relevant to the assessment endpoint (e.g., survival, growth and reproduction of piscivorous mammals). Effects that are transient or occur only at lower levels of organization (e.g., biochemical effects that are not translated and/or related to effects to higher levels of organization) are not typically considered “ecologically significant” effects.

9. The use of the HQ approach and the definition of what constitutes “low”, “moderate”, and “high” risk, either quantitatively or qualitatively should be formally and fully articulated in the ERA. In some cases an HQ of less than 1 is considered indicative of no risk, but this designation is not applied consistently in the ERA.

RESPONSE O-RS-19

Please refer to the response to General Issue 7.

10. The exposure estimates of benthic invertebrates to tPCBs and TEQ in the laboratory-based sediment toxicity studies should be revised and based on the most synoptic sediment chemistry data. The current exposure estimate and the subsequent MATC of 3 mg/kg currently is not reflective of the most synoptic sediment chemistry data, and thereby, is inappropriate.

RESPONSE O-RS-20

Please refer to the response to General Issue 6.C.

11. Because the results of the wood frog studies by FEL and GE are the most comprehensive and potentially the primary basis for estimating potential risks to amphibians inhabiting the Housatonic River watershed, I recommend that EPA re-evaluate the Resetarits study, and where feasible, incorporate those data into the modeling projections conducted by Dr. Ferson.

RESPONSE O-RS-21

EPA gave careful consideration to the Resetarits report during the preparation of the ERA. There were a number of limitations to the study that constrained the use of the data, as described in the ERA and summarized below:

The description of the study methods lacked any details on the collection of eggs, laboratory conditions where the eggs and larvae were kept (e.g., vessel size, water chemistry, temperature), how the larvae were assigned to each treatment, observations of mortality or abnormalities, observations made in the field, atmospheric and climatic conditions, natural predation, availability of food in enclosures, opportunities for metamorphs escaping, analytical methods for lipids and PCBs, and sediment sampling protocol.

Recent literature on amphibian toxicity has identified several important considerations for the establishment of sensitive toxicity tests. First, the physiological changes that occur in amphibians must be taken into account in the design of amphibian toxicity experiments, in particular, the biological changes (e.g., surge in thyroid hormones) associated with the metamorphosis. Consequently, consideration of the sensitivity of various life stages is crucial to the assessment of organism sensitivity to COPCs. The GE study focused on only a portion of the amphibian life cycle and did not present a rationale for this choice. Additionally, investigations were not carried through the metamorphic period for all test animals; metamorphosis has been shown to be a period of particularly high sensitivity to contaminant exposure.

Therefore, while the procedures used by Resetarits were relatively sound, the design of the study did not include the necessary measurements and/or life stages to inform the risk assessment.

The study was not able to detect effects representative of field exposures to amphibians in the Housatonic River floodplain, by failing to simulate the exposure conditions relevant to wood frogs (i.e., sediment exposures of 1 to 100 mg/kg through to metamorphosis). Wood frog larvae were placed in relatively clean sediment throughout the experimental period. The experimental animals were placed in two vernal pools (23b-VP-1 and 23b-VP-2), with sediment tPCB concentrations of 0.21 and 0.3 mg/kg (Appendix E, Table E.2-8), respectively. Therefore, exposure of developing larvae to PCBs was not representative of that occurring in the field, except in the cleanest pools in the floodplain.

Vernal pool sediment and water tPCB concentrations have been shown to be a significant factor in the bioaccumulation of PCBs and the subsequent manifestation of effects as frogs matured in the pools (FEL, 2002). Of the 68 vernal pools evaluated for sediment PCB concentration (during various investigations), only 6 had tPCB concentrations < 1 mg/kg. Thus, approximately 91% of the vernal pools evaluated within the PSA had sediment tPCB concentrations of >1 mg/kg. The Resetarits study did not expose the developing larvae to the range of sediment tPCB concentrations that are characteristic of the wood frog breeding pools in the Housatonic River floodplain. This assumption in the experimental design, therefore, biases the study against the detection of effects.

12. Based on the January 2004 public meetings, there appears to be a substantial information base on DELTs in local finfish populations. If accurate, I recommend that this information be more fully displayed and discussed in the main text of the ERA. It is potentially a key element of the field work that does not appear to have been highlighted in the ERA.

RESPONSE O-RS-22

Please refer to the response to General Issue 13.E.

13. Because it has potential implications for the estimate of exposure (and risk) for piscivorous birds and mammals, I recommend that the tissue analyses conducted for any finfish in the Housatonic River be fully displayed and discussed in the main text of the ERA. These data appear to exist but are not readily found in the ERA or the appendices. The existing datasets should not be truncated so that only tPCBs and TEQ are provided.

RESPONSE O-RS-23

Please refer to the response to General Issue 1.

14. The estimate of risk to piscivorous birds is uncertain and may warrant further evaluation by EPA. The lack of a strong field study, coupled with the results of the modeled exposure and effects suggests this assessment endpoint may not have been adequately evaluated in the ERA. This does not negate the findings of the tree swallow study. However, given the differences in feeding between insectivorous and piscivorous birds,

one cannot conclude that the results in tree swallows is or could be applicable to belted kingfishers, osprey or other piscivorous birds. Therefore, I recommend that EPA collect and evaluate biological survey data on populations of osprey that inhabit the Housatonic River watershed to help reduce the uncertainty associated with potential risk to this species. If there are no biological survey data, or, as GE has contended, osprey do not inhabit the Housatonic River watershed, then EPA should consider another piscivorous bird species for the conduct of the ERA. Where biological survey data are available for heron in the Housatonic River watershed, then it is reasonable to pursue this receptor in the ERA in lieu of the osprey.

RESPONSE O-RS-24

EPA concurs that the results of the tree swallow study should not be used to estimate risks to piscivorous birds. No such extrapolation was done in the current ERA, nor will it be done in the revised ERA.

EPA believes that osprey are an appropriate representative species for piscivorous birds because:

- There is a high likelihood that ospreys will be nesting in the PSA in the near future as they continue to recover from exposure to DDT and other organochlorine pesticides.
- Ospreys have already occupied appropriate coastal area habitats and are now moving inland in CT and MA.
- The PSA has suitable habitat for ospreys.
- During the EPA field studies, a pair of ospreys in the PSA displayed courtship behavior during the mating season, but failed to nest.

EPA does not consider great blue heron to be an appropriate representative species for piscivorous birds because:

- There are likely only 3 to 5 herons from the rookery that forage in the PSA above Woods Pond Dam.
- No reproductive data are available for these particular birds because it is impossible to know which nests in the rookery belong to the herons that forage in the PSA.
- Reproductive data are only available for the entire rookery which consists of approximately 28 nests. More than 80% of these birds do not forage in the PSA above Woods Pond Dam because herons are territorial feeders.
- Thus, any impacts on individuals that forage in PSA above Woods Pond Dam would be obscured by the greater than 80% of birds that are not being exposed to PCBs.

15. There is uncertainty associated with the cause of death of kits in the mink feeding study. It appears that kits which died unexpectedly in the study were not necropsied nor the cause of death determined. If however, these animals were necropsied and the cause of death determined, I recommend that information be provided in the ERA and clearly

discussed. Otherwise, there doubts will remain as to whether the drop in kit survival at 6 mos of age is due to maternal exposure to tPCBs or due to other causes.

RESPONSE O-RS-25

All kits that died after 6 weeks of age were necropsied; the pathologist's assessment is contained in the mink feeding study report (all died of infections thought to be unrelated to PCBs) (Bursian et al 2003). Kits that die prior to weaning at 6 weeks are not necropsied as part of the standard operating procedure used by MSU for the following reasons. Kits are handled only 3 times during this period (to obtain body weights at birth and at 3 and 6 weeks of age). Excessive disruption of the female can cause her to eat her young, particularly from birth to 3 weeks of age. If a kit is unhealthy or if it dies, the mother will often eat it, leaving no carcass to necropsy. If a kit dies and is not disposed of by the female, it usually is not a candidate for necropsy because of the length of time between death and discovery of the carcass, which often is buried in the bedding within the nest box. In addition to the mother cannibalizing their young, the kits may cannibalize a weaker sibling as they approach six weeks of age. Therefore, EPA does not agree with the Reviewer's comments that the study should be discarded because it was not possible to perform necropsies on kits less than 6 weeks of age.

The ERA includes a discussion of the body weights of kits. Those in the high dose group at 3 weeks of age had significantly lower body weights than the controls, which implies a treatment-related effect on growth (a known effect of PCB toxicity to mink). This was followed by significant mortality from 3 to 6 weeks of age in the same dose group (as well as a return of average body weights in this treatment to control values), which the Principal Investigator interprets as those kits that were low in body weight at 3 weeks of age died during the subsequent 3-week period. Previous mink studies have shown kit survivability to be a sensitive indicator of PCB toxicity in this species and at least one study has demonstrated significant delivery of PCBs from female to kit during lactation (Bleavins et al. 1980; Heaton et al. 1995). Neither of these studies necropsied kits that died prior to 6 weeks of age. More importantly, deaths due to causes other than contaminants in the diet would be expected to occur randomly across all dietary treatments. This was not the case, as there was a significant effect on growth of mink kits from 0 to 3 weeks of age, and a significant effect on survival of mink kits from 3 to 6 weeks of age in the highest dose treatment.

16. There was significant debate during the January 2004 public meetings as to the statistical analysis for the Boonstra study on small mammals, and this was reflected within the ERA. Because the conclusions drawn for small mammals are important for determining the need for potential remediation of soils in the PSA, it is imperative that one and only one mutually acceptable approach is presented in the ERA. I recommend that EPA and GE select a single statistical approach and provide that one approach and the results thereof in the ERA.

RESPONSE O-RS-26

Please refer to the response to General Issue 16.B.

17. In my opinion, the use of rodent data to support a modeled estimate of effects in red fox is not scientifically supportable. The uncertainty is so large that there is little to be gained from this evaluation. I recommend that EPA delete the red fox assessment endpoint from this ERA. There are sufficient data from the other assessment endpoints on which to base a final risk management decision.

RESPONSE O-RS-27

A detailed review of the toxicity of tPCBs and TEQ to mammals, including toxicity to red fox, was conducted. Toxicity studies on red fox and close surrogates (e.g., canine species) were not available for tPCBs and TEQ. A summary of the toxicity of tPCBs and TEQ to mammals is presented in Appendix J.3. EPA agrees that there is uncertainty when extrapolating effects data for rodents to red fox. This uncertainty was acknowledged in the Sources of Uncertainty section (Section J.4.5) -- "The largest source of uncertainty in the effects assessment was associated with the lack of toxicity studies involving the representative species." Moreover, Section J.4.5 states that, "These extrapolations introduced uncertainty in the effects assessment because of the variations in physiological and biochemical differences between the species such as uptake, metabolism, and disposition that can alter the potential toxicity of a contaminant." EPA will expand the discussion surrounding the extrapolation of rodent effects data to red fox, and revisit the literature to see if additional toxicity data are now available.

18. The Land H-statistic discussion in the main text of the ERA and in some of the Appendices is complex and not easily followed. I recommend that the authors consider a re-writing of the description to simplify the strengths and weaknesses of the approach.

RESPONSE O-RS-28

EPA will reexamine the text describing how the Land H-statistic was used in the ecological risk assessment and, if possible, simplify the discussion. However, EPA believes it is important to specify the methods and decision rules used in applying the Land H-statistic, as done in Section 6.3. The decision rules are complex, but for reasons of transparency, they are listed in their entirety in this section.

19. I recommend that the biological surveys conducted for the T&E species be included in the weight-of-evidence evaluation. Given the magnitude of the risk question at hand it is important that all relevant information be brought to bear in the evaluation of potential risk. To exclude this information on the basis that it is not quantitative is not appropriate, objective or reasonable.

RESPONSE O-RS-29

The biological surveys for T&E species were included in the weight-of-evidence evaluation. In the revised ERA, EPA will describe in more detail how field surveys were incorporated into the weight-of-evidence evaluation beyond the information presented in Table 11.4-4.

Timothy Thompson:

This memo represents my final comments to the ERA conducted for the Housatonic River as part of the Peer Review Panel Process. This document is modified from the pre-meeting comments submitted January 5, 2004, based upon the review and discussion from the Peer Review Panel meeting held January 14 – 16.

The format here follows that prescribed in the Charge to the Peer Reviewers provided by SRA International Inc. I have deviated somewhat from that format by answering each of the Charge Questions as they relate to the Assessment Endpoints identified as part of Charge Question 3.

I would like to reiterate what I have said at our Peer Review meetings, and in my January 5 comments – this is an amazing document with a considerable amount of thought and effort expended by both EPA and GE scientists and ecologists. I extend my complements to each of them for the tremendous effort – both in terms of the science they practiced and the Herculean effort needed to bring these documents together. I also reiterate that it has been my pleasure to observe the very professional way in which EPA and GE interact on this project. Reasonable people can, and will disagree. It is very nice to observe the process do so in a reasonable way.

An explanation to the way I approached my comments to the ERA may be helpful. I agree strongly with the following statement made in Section 2.8.1 (page 2-59): "Ultimately, the value of an ERA depends on whether it can be used to determine if a baseline risk is present and to support managerial decisions." As such, I have tried to "truth test" as many of the conclusions as was possible within the time frame allotted. This meant evaluating in detail the Supplemental Investigation Work Plan for the ERA, requesting the same – and receiving work plans from GE (thank you!), reading through the RCRA Facilities Investigation, as well as the ERA, supporting Appendices, and all of the other supporting documents. Where I have made comments, it is in the interest of pointing out what was apparently intended (work plans), what was actually done (appendices and supporting documents), and how that was used in the characterization of risk.

I would like to acknowledge the remarkable diversity and intellect of my Peer Review colleagues. This was truly a powerful assemblage of scientist who critically, and objectively evaluated the studies conducted by both EPA and GE with no bias – save for what their experience and intellect told them was right and wrong. I was humbled to be considered amongst such a group, and believe that the ERA can only be strengthened by their individual, and our collective input.

1. CHARACTERIZATION OF HOUSATONIC RIVER ECOSYSTEM

In evaluating the general items specified in the Consent Decree listed above, the Peer Review Panel members shall give specific consideration to the following questions:

1. ***Was the ecosystem of the Housatonic River watershed properly characterized, and was this information appropriately applied in the Problem Formulation and subsequently in the ERA?***

Valery Forbes:

My answers to the Charge questions are based primarily on the main ERA but include, where relevant, EPA's responses to Panelists' written questions and oral responses provided at the public meeting held 13-16 January 2004. Thus I am assuming that if the requested information was not present in the main ERA but was addressed satisfactorily in the EPA's written or oral responses that appropriate amendments will be made following the Peer Review meeting.

RESPONSE 1-VF-1

EPA will revise the ERA as specified in responses to the written and oral comments from the Reviewers.

Comments: The ecological characterization seems to have been extremely thorough, and a relatively detailed knowledge of the ecology and habitat usage, particularly of the birds and mammals, seems to have been incorporated into the ERA. However I feel it is unsatisfactory that the assessment endpoints were chosen, to some extent, on the basis of whether or not data were available for the species under consideration (EPA response to Panel Question BS1). I would argue that the availability of data is not an appropriate criterion for selection of assessment endpoints (though it can be a constraint for selecting measurement endpoints). If there is an endpoint for which protection is deemed an appropriate goal on the basis of the site characterization, then the necessary data should be collected as part of the ERA.

RESPONSE 1-VF-2

Please refer to the response to General Issue 2.

Proposed Changes: A detailed road map or data inventory could increase clarity and reader-friendliness. A figure (or series of figures) showing spatial variation of tissue sample sites and concentrations could be a useful addition.

RESPONSE 1-VF-3

Please refer to the response to General Issue 1.

The ERA should include an explanation of why some of the risk characterization studies were not included in the ERA (e.g., dragon flies, mussel, blue gills). Better overviews (tables or figures) of what data have been used would improve the document.

RESPONSE 1-VF-4

Please refer to the responses to General Issues 1 and 4.

Thomas W. La Point:

There is no doubt that much effort and thought have been undertaken by the U.S. EPA and GE. I applaud them and congratulate them on the tremendous amount of research, thought and effort that went into this ERA. It is always easy to "Monday-morning quarterback" and criticize the resulting documents. I hope my criticisms are not taken in this manner. They are meant to suggest my concerns and perhaps point towards some "confirmatory" sampling that may lessen some of the uncertainty associated with various assessment or measurement endpoints.

The Housatonic River ecosystem ("Ecosystem") was very well characterized. The report by Woodlot (2003) provides an excellent overview of the variety of studies conducted and species sampled or identified. However, it is unfortunate that neither bats nor waterfowl (examples: mergansers, fish-eating ducks or "dabbling" ducks that filter invertebrates from sediments) were included in the Risk Assessment Conceptual model (diagrammed in Fig 2.7-1 or included as assessment or measurement endpoints (Table 2.8-1).

RESPONSE 1-TL-1

Although bats were not included in the conceptual model diagram, they were included in the risk assessment. The small-footed myotis was one of three species chosen for assessment of effects to rare, threatened, and endangered species. Appendix K of the ERA contains details of the risk characterization for this bat species.

In response to Reviewer's comments, EPA will add a waterfowl species to the revised ERA as an additional representative species.

Also, given the importance to this ERA in assessing the risk to terrestrial species, including amphibians and salamanders, that less effort was devoted to sampling soil invertebrates such as earthworms, isopods, beetle larvae, millipedes and centipedes. Given that these were sampled in 2000 (Woodlot Report), and given the numbers of isopods, cicadas and slugs sampled, these may well be important routes of exposure for predators.

RESPONSE 1-TL-2

EPA concurs that terrestrial invertebrates are a route of exposure for predators feeding in the floodplain; however, the number of terrestrial invertebrate samples was limited, in part, because better data existed. Predator tissue samples were collected for robin eggs, nestlings, and for several species of small mammals, all of which feed in the floodplain. This provided direct measures of tPCB concentrations for an insectivorous bird and for herbivorous, omnivorous, and insectivorous small mammals. Direct measurement introduces less uncertainty in the estimate of contaminant uptake than exposure modeling.

Finally, I am concerned that there is no comprehensive overview or integration of the species in the ecosystem mapped out in Figure 2.7-1. Given the conclusions (see more on this below) of "high risk" to benthic invertebrates, that information is not carried into the risk assessment for insectivorous birds. The birds are judged on strictly other, limited information. However, if their food base is judged to be at high risk, then the predatory organisms relying on this food base must also be judged to be at risk. In any case, there appears to have been little-to-no overview linking the variety of seemingly-independent assessment endpoints. After all, this is an ecological risk assessment – and one must view the resident species as an interacting biological community.

RESPONSE 1-TL-3

EPA agrees that indirect effects, such as effect on food supply, can have an influence on population dynamics. A qualitative discussion of how indirect effects can impact population and community dynamics was presented in ERA Section 12.4.2.2.

These ideas were discussed in another context during the Lenox public meeting. In my opinion, the EPA should review the prior documents detailing the goals of this ERA. An earlier memo states that the goal is to "ensure recovery and maintenance of healthy local populations and communities of biota." The question is: "how was this defined for the purposes of the risk Housatonic Rest of River assessment?

RESPONSE 1-TL-4

Please refer to the responses to General Issues 2 and 9.

James T. Oris:

Some of the best work done in the ERA was the ecological characterization portions. The document describing the ecosystem of the Housatonic River watershed is very thorough and clear with few exceptions.

<u>Comment 1.1. Appendix A.1, section 1.1</u>: Purpose of document states 8 work plans were addressed. Of these workplans listed, there is no mention of fish, macroinvertebrates (other than dragonflies), or soil infauna. Yet, there are entire chapters devoted to each of these groups in the document. Further explanation at this point in the document would clarify the role of the groups not included in the listed workplans.

RESPONSE 1-JO-1

Section 3.0 of Appendix A.1 discusses how fish, macroinvertebrates, and soil infauna data were incorporated into the ecological characterization. The ecological characterization was a compilation of data collected both prior to and during the ERA. Detailed descriptions for each set of data can be found in Sections 3.1 through 3.2.6.

<u>Comment 1.2. Appendix A.1, Attachment C, Species matrix</u>. What does "SC" mean in species status? Also, there is no analysis of this data (at least for fish) for community quality. For example, for the fish community, Attachment C indicates that 49% of the suitable habitat in PSA was not occupied with expected fish species. I believe a fish community analysis would be appropriate beyond just the description of what is there.

RESPONSE 1-JO-2

"SC" refers to a species of "special concern," a Commonwealth of Massachusetts designation accorded to certain species considered to have suffered a decline that could threaten the species if allowed to continue unchecked, or which occur in small numbers or with such restricted distribution or specialized habitat requirements that they could easily become threatened in MA. This and other definitions are described in Table 1-1, Section III-1-2 of Appendix A.1.

In Appendix A.1, Section III, Chapter 3 presented a detailed narrative description of the fish community found in each reach of the PSA. Both qualitative and quantitative data on species richness, abundance (i.e., biomass), and species:habitat associations are included. Attachment C is a species:habitat matrix table that lists species occurrence per habitat. The data in the table indicate that suitable habitat in the PSA was occupied by the expected fish species.

Comment 1.3. It is not clear whether all the studies in the ecological characterization were utilized in the ERA. Specifically, the dragonfly data (which is quite robust) is apparently not used in the ERA.

RESPONSE 1-JO-3

Please refer to the response to General Issue 4.

Comment 1.4. An inventory of data collected and used for both abiotic and biotic characterization is necessary to clarify what is present in the document. This will help in the presentation of the document and help clarify the decision processes used in the problem formulation phase.

RESPONSE 1-JO-4

Please refer to the response to General Issue 1.

Comment 1.5. The Massachusetts portions of the river have been extensively characterized, but the Connecticut reaches have only been treated superficially. This needs to be justified in the document, or further characterization needs to be done.

RESPONSE 1-JO-5

Please refer to the response to General Issue 5.

Mary Ann Ottinger:

General Comments:

The Pre-ERA provided an appropriate characterization of the ecosystem of the Housatonic River watershed. The surveys of wildlife, benthic and aquatic communities, and overview of the habitats provided a thorough review of the constituents of the ecosystem and the elements of the ecosystem that warranted further assessment. The characterization included data collected most intensively over the last 10 years, with some information available from earlier studies.

Unfortunately, there appear to be few historical data available on species richness in the region at the time of early stages of contamination by chemicals of concern (COCs). Therefore the pre-ERA assessment and species surveys represent populations that potentially have already been impacted by the COCs. This is relevant when considering reference areas, especially in the case of wildlife that may move within the region over successive generations from an area relatively unaffected by the COCs to the primary study area (PSA) or vice versa.

RESPONSE 1-MO-1

EPA concurs that reference areas, particularly if close to the PSA, provide a context for evaluating species richness in the absence of contamination. EPA also recognizes that quantitative data on species occurrence, distribution, and richness do not exist for the PSA prior to releases of PCBs. Such data are very uncommon for most places in MA, except for a few National Wildlife Refuges or similarly protected lands.

In light of the available surveys and information, the problem formulation was logically constructed and focused on the relevant classes of organisms. In the case of avian species, several have been considered as representative species in the categories of insectivorous birds, piscivorous birds, and threatened and endangered species. However, the choice of the representative avian species did not include ground dwelling species, such as turkeys, quail, or pheasants, which would receive exposure through seeds and sediment.

RESPONSE 1-MO-2

Due to practical limitations and the likelihood that the assessment would have quite a bit of uncertainty due to absence of site-specific studies and species-specific toxicity data, EPA will not be adding a ground-dwelling bird species to the revised ERA.

Moreover, turkey eggs and mallard ducks, which were collected opportunistically, had measurable tPCBs, indicating exposure of these species.

RESPONSE 1-MO-3

Turkey and mallard eggs were not collected by EPA (EPA did conduct a sampling and analysis of wood duck eggs in 2004), nor is EPA aware of any egg tissue data for turkey or mallards from the PSA. Both mallard and wood ducks were collected by EPA, and breast and liver tissue were analyzed.

Similarly, the woodcock takes in approximately 10% (dry weight of intake) sediment in its diet (Rattner et al., Handbook of Ecotoxicology, p. 157; 2003). Although known to occur in the PSA, the woodcock was not selected as a representative species because it is too secretive and therefore difficult to study. This is reasonable, but the choice of the kingfisher as a representative species and the continuation of the study when difficulties arose becomes questionable. Therefore, it is not clear that the choice of representative wildlife species was appropriate in all cases.

RESPONSE 1-MO-4

See Response 1-MO-2. Please note that GE chose to perform the kingfisher study without concurrence or input from EPA. EPA intends to add a duck species to the revised ERA.

Specific Comments (references to the ERA documents are included in some sections):

[1.ES-4] The ERA focus is on "the river from the confluence, 2 miles below the GE facility to Woods Pond Dam; the pond is the first impoundment downstream, with a depositional environment". Ecological characterization was conducted in detail and some of the GE studies are incorporated. It is not clear if these earlier characterizations were thorough. The inventory conducted as part of the ecological characterization included reference areas. A mention was made of relatively high calcium levels in the Housatonic River, which was not the case in many of the reference areas. Therefore, reference areas have additional variables that would potentially confound comparisons.

RESPONSE 1-MO-5

EPA concurs that reference areas may result in some additional variables that can complicate certain comparisons; however, EPA believes that the ecological characterizations that were performed were extremely thorough. During the development of the risk assessment, EPA gave careful consideration to the potential confounding factors associated with the selection of reference areas. Appendix A.1 (Section II, pages 15-17) contains descriptions of reference areas and their underlying geology. Two of the four reference areas, Hinsdale Flats WMA and Threemile Pond WMA, both have bedrock and high calcium level characteristics similar to those found in the PSA.

[1.ES-12] Risk characterization includes site-specific studies, result analysis, conclusions, and Weight of Evidence (WOE), with consideration of uncertainties for the endpoint and risk to receptors (outside representative species) and risk to downstream receptors. This is an appropriate and logical sequence. Playback data yielded assurance that the species were present, but these were not necessarily the species that were examined as part of the ERA. Further, there was little opportunity for following up on the surveys to ascertain the species diversity over several years and the historical data are not complete relative to species surveys. Therefore, the characterization included a record of species observed and their habitat, but were insufficient for assessing populations and consequently for assessing impact to populations.

RESPONSE 1-MO-6

Playback call data were used to identify species that occur in the PSA, as part of the ecological characterization - the focus was not solely on the representative species selected for the ERA. As discussed in the response for General Issue 4, representative species for detailed analysis of risks were chosen using information gathered in the ecological characterization. A comparison was then made (see Section 12.4 of the ERA) between representative species and similar species that were not explicitly considered in the quantitative ecological risk assessments. The ecological characterization was not intended to assess populations or impacts to populations.

[I. B6] Clarify the description of soil background levels ("not detected at a sample quantitation limit of less than 0.5mg/kg or detected at a concentration of less than 0.3 mg/kg). As stated in later in the document, it appears that measurements below assay sensitivity are assigned ½ the level between zero and the sensitivity limit. If this is correct, then please clarify in this section.

RESPONSE 1-MO-7

The objective of the soil background discussion on page B–6 of the Pre-ERA was to provide a detailed description of the criteria used to identify soil samples within the PSA that could serve as background samples for subsequent screening of inorganic COPCs. The PCB quantitation limits discussed were presented to illustrate that all samples selected were not influenced by site related contamination. PCBs were not included in the soil COPC screening against background levels; they were used only to help identify unimpacted soil samples collected within the PSA.

[B16-Soil Preliminary Remediation Goals (PRG)] Lowest observed adverse effect level (LOAEL) from Sample et al., 1996 for soil was based on the woodcock and the shrew; however the woodcock was not evaluated in the ERA. The rationale given was that the woodcock is secretive and not easy to find. Conversely, several avian species that were used for modeling also were not found in the PSA and the kingfisher that was studied occurred in such low numbers that there was low confidence in the results of the study.

RESPONSE 1-MO-8

In the Pre-ERA, toxicological benchmarks were developed for water, soil, sediment and tissue for each contaminant of potential concern (COPC). These benchmarks were intended to be conservative. Thus, when concentrations of a COPC in the PSA were below the corresponding benchmark, the COPC was assumed to pose negligible risk to biota for the medium under consideration. In these cases, the COPC was not considered any further in the ERA.

The American woodcock (*Scolopax minor*), a ground-dwelling species, has a high potential for exposure to soil contaminants relative to most other avian species (Efroymson et al. 1997). Therefore, soil benchmarks for birds were based on woodcock exposure (see Appendix B). As a result, the soil benchmarks for birds are conservative and should be protective of other birds. EPA believes that this is the appropriate approach for a contaminant screening process.

[B18] No observed adverse effect level (NOAEL) was determined from "effects of potential "ecological significance" evaluated (e.g., lethality and reproductive effects)". As will be discussed below, these endpoints are not necessarily sensitive to the COCs in some of the representative species. In addition, aside from embryos and prepubescent animals, there was little consideration of age-related or seasonal effects in the ERA as they impact some species.

RESPONSE 1-MO-9

While EPA agrees that lethality and reproductive effects might not be the most sensitive endpoints of ecological significance, the TRVs selected for the Pre-ERA were based on the most conservative toxicological endpoints identified after an extensive review of the scientific literature. EPA, therefore, believes the TRVs used to screen COPCs were conservative and appropriate.

[B25] Tier III elimination of PAHs and pesticides was based on the presence of these compounds in fish tissue or on likelihood of metabolism in fish. This eliminates consideration of these compounds for other species and may be a confounding factor due to mixture effects.

RESPONSE 1-MO-10

EPA assumes this comment refers to the potential for PAHs and pesticides to biomagnify in the aquatic food chain.

PAHs were not included as part of the fish tissue analysis on the Housatonic River because of their low potential to biomagnify and the ability of fish to metabolize most PAH compounds (Eisler 1987). Therefore, PAHs were not considered when evaluating exposure to contaminants through the fish ingestion pathway. It should be noted that several PAHs were retained as sediment COPCs after the Tier III evaluation and were further evaluated as part of the fish assessment endpoint analysis.

Several pesticides were retained as COPCs in fish tissue following the Tier III evaluation. Pesticides in fish tissue were eliminated as COPCs for specific piscivorous receptors after a thorough review of the fish tissue re-analysis (see Attachment B.2). Justification for the elimination of pesticides as COPCs in fish is presented in Section 2.4.

[Attachment B.2] Reanalysis of pesticides in fish tissue (only) resulted in lowered baseline risk assessment. How does this impact wildlife and what are the implications for the validity of the measurements?

RESPONSE 1-MO-11

As a result of the fish tissue re-analysis (see Attachment B.2) and the resulting implications for concentrations of pesticides identified in fish tissue, EPA believes the potential for ecological harm to piscivorous receptors (ingesting pesticide-contaminated fish) is low relative to the risks associated with tPCBs and dioxin-like TEQ. As indicated in Table B-147 of the Pre-ERA, pesticides in soil were not retained as COPCs for wildlife following a conservative screening process.

Brad Sample:

In terms of characterization of contamination in abiotic media, Reaches 5 and 6 in the Housatonic River watershed appear to be very well characterized. Abiotic characterization of downstream reaches is much less extensive. In contrast to the abiotic media, characterization of contaminants in all biota types is not as robust. The number of biota samples and the spatial distribution of these samples is more limited. Further, because maps presenting biota sampling locations to concentrations in abiotic media are not presented, it is difficult to determine whether the biota chemical data adequately represent the full range of concentrations addressed by the abiotic data. It is essential that biota concentration data represent the locations with maximum COPEC concentrations in abiotic media.

RESPONSE 1-BS-1

Please refer to the response to General Issue 4.

Section 2.5 - Although I recognize that PCBs are the primary COPC for the site, details on the distribution of other COPCs should also be presented. If not in this section then at least in an appendix. At a minimum, dioxins/furans need to be included in this section as they are the second-most dominant COPC. It is important to know how TEQs vary in relation to PCBs. Another issue - although graphical presentation of data is very useful, tabular presentation of the data (summary statistics by reach) plus information about the datasets on which they are based (when where they collected, by whom, sampling methods used, sample handling procedures, etc.) should be provided. It is very difficult to track the data sources and summaries for this assessment.

RESPONSE 1-BS-2

Please refer to the responses to General Issues 1 and 3.

Figures 2.5-5 through 2.5-11: These figures are very useful for displaying the spatial distribution of PCBs in floodplain soils. Similar figures that displayed PCBs in sediments and TEQs in floodplain soils and sediments would aid the understanding of the spatial context of contamination. An additional enhancement would be to include notations as to where biological sampling was conducted - key maps would display locations where biota sampling for tissue analyses was performed. Additional maps that showed locations where field studies were conducted (i.e., swallow nest boxes, kingfisher burrows, robin nests, small mammal population surveys, etc.) in relation to soil/sediment concentrations would greatly enhance the utility of the report.

RESPONSE 1-BS-3

Please refer to the response to General Issue 1.

Much more detailed information on the biota chemistry data are needed. Details on where all samples were collected cannot be readily located in the report. Details on how all samples were collected and processed, which analytes were measured in each, and whether samples of abiotic media were collected at the same locations are not readily apparent in the report. Because this report relies heavily on these data, much more detailed information on the chemical characterization of biota is needed.

RESPONSE 1-BS-4

Please refer to the response to General Issue 1.

Report is not transparent as to which data were used and where they reside in the report. A road map or data inventory that tracks all data that contributed to the decision for each assessment endpoint is needed. The data inventory should include where the data may be found in the report, what they are comprised of (i.e., field analytical chemistry data versus field survey data versus laboratory bioassay data), and the spatial location from which they were obtained and therefore represent. Inclusion of something like this in the beginning of each section would allow readers to develop a better understanding of the depth, robustness, spatial representativeness of the data upon which the risk conclusions for each endpoint are based.

RESPONSE 1-BS-5

Please refer to the response to General Issue 1.

Characterization of the Connecticut section of the river does not appear adequate – biota sampling and sediment sampling much more limited. I recommend that EPA consider more sampling in this area, especially in depositional areas behind the multiple dams that occur here.

RESPONSE 1-BS-6

Please refer to the response to General Issue 5.

Ralph G. Stahl:

I believe the ecosystem of the Housatonic River was properly characterized within the boundaries of Massachusetts, but it is not clear that the same rigor was applied to sections of the watershed in Connecticut. I found the comments of Dr. de Fur at the January public meeting to be persuasive with regards to the need for additional sediment characterization in the Housatonic River watershed in Connecticut. In this regard, the characterization of the ecosystem of the Housatonic River watershed was not consistent. I recommend that EPA consider additional characterization of the sediments located behind dams in Connecticut that are in closest proximity to the PSA. It is likely that, over time, sediments and their PCBs have been transported downstream and have settled behind one or more of these structures. Thus while the concentrations of PCBs decline downstream of the PSA, as would be expected in the absence of a continued source of the PCBs, this decline would not be observed in sediments deposited behind downstream dams. Uncertainty remains high as to whether or not PCBs, at potentially problematic concentrations, are present in these depositional areas and simply have not been characterized. Characterization of sediments is not needed for all dams on the Housatonic River in Connecticut, but should be sufficient to reassure the public and decision makers that this issue has not been overlooked.

RESPONSE 1-RS-1

Please refer to the response to General Issue 5.

I do not believe that additional characterization is needed for soils in the floodplain within Connecticut; however, additional floodplain sampling may be beneficial in areas immediately downstream of the PSA in Massachusetts. These additional data may be helpful in refining some of the exposure estimates for terrestrial receptors such as birds, mammals and amphibians. As stated previously, it would be very beneficial to develop detailed maps which illustrate the synoptic sampling of floodplain soils, sediments, water column, and biological tissues. Where possible, results from this sampling should be displayed for each of the reaches of the river and the associated reference locations.

RESPONSE 1-RS-2

Several hundred samples were collected from the floodplain below the PSA in MA; EPA believes this amount of data is sufficient to adequately estimate risk in this area. With regard to maps, please refer to the response to General Issue 1.

Timothy Thompson:

Benthic Invertebrates

There is a tremendous amount of ecosystem information that has been generated that could be used to help articulate the importance of aquatic invertebrates to the Housatonic River. Numerous studies have been conducted related to aquatic insects with the Housatonic River, and in the PSA. Given that the basis for evaluating aquatic invertebrates is the importance in the food chain and transfer of PCBs into higher organisms, a discussion of the higher trophic level organisms and the prey upon which they depend should be better articulated for this assessment endpoint.

RESPONSE 1-TT-1

EPA concurs that aquatic invertebrates play an important role in transferring PCBs into higher trophic-level organisms. Table 12.4-1 provides details on the comparative risks of PCBs to representative and other species in the Housatonic River based on potential exposure, including diet. Aquatic invertebrates are an exposure pathway for multiple species of fish, amphibians, and birds. In the revised ERA the EPA will describe in greater detail the role of aquatic invertebrates in the food chain and the transfer of PCBs.

One consideration is to make more use of the riverine classifications in describing the various benthic habitats, expected species, and anticipated routes of exposure. Within the Natural Communities descriptions, for medium gradient streams the soil type is classified as cobble, gravel and sand, with moderate stream flows. Within the classification it would then have been useful to list the type of aquatic invertebrates expected in those habitats – and potentially have actually quantified the species existing there (presumably higher in epibenthic organisms, and in particular EPT species). Aquatic invertebrates in this habitat would be less likely to be exposed through fine-sediment contact and porewater uptake, and would receive PCBs through the water column and/or contaminated food. Within depositional pools, such as the moderately alkaline pond, more benthic infauna would be encountered, and would include a significant component of the invertebrate biomass exposed to PCBs through sediment and porewater ingestion. But in addition, given the high density of emergent and floating vegetation, a high degree of "benthic" invertebrate production would also be expected from within the plant mass.

RESPONSE 1-TT-2

The EPA benthic community sampling program emphasized substrates that are representative of the majority of sediment located within the PSA. EPA concurs that the macrophyte community in the river, backwaters, and Woods Pond also provides habitat for a variety of aquatic invertebrates, and that invertebrates exposed primarily via water column pathways are not well represented by the benthic grab samples. There are also localized areas of sediment substrate (i.e., cobble/riffle habitat) that comprise a small percentage of the PSA habitat and that have different assemblages of organisms.

EPA will provide more information on the types of aquatic invertebrates expected in the revised benthic invertebrates ERA. This will include a description of the ranges of substrate types observed in the PSA and the linkage between habitat and organism assemblages. Information on assemblages found in coarser habitats is available in Chadwick (1994), and will be integrated into the ERA.

The conceptual model of the PSA assigns invertebrates that obtain their PCB exposures through the water column to a category called "water column invertebrates." This category includes epiphytic crustaceans on macrophytes or other substrates that allow the organism to be physically isolated from sediment and pore water. As such, these organisms are not considered "benthic invertebrates." The concentrations in water column invertebrates are lower than benthic invertebrates (based on bioaccumulation model predictions), so the assessment based on benthic invertebrates is also expected to be protective of these organisms.

A general comment here, and throughout the ERA, is that figures showing spatial variation of sample sites, including sediments, soils, and tissues should be included to help illustrate the depth and breadth of coverage.

RESPONSE 1-TT-3

Please refer to the response to General Issue 1.

The ERA also needs to make a better explanation of the types of communities encountered, relative to the grain size and organic carbon contents in those sediments. As noted by Dr. Oris in his pre-meeting comments [3.1(c).1], the relationships between chiromid abundance, taxonomic diversity relative to substrate (fine vs. coarse) potentially obfuscate any relationship between tPCBs and community structure. Without a complete discussion of communities, substrates, riverine conditions, and organic carbon, the relationship to tPCBs and toxicity is tenuous.

RESPONSE 1-TT-4

Please refer to the responses to General Issues 11.A, 11.D, and 11.G.

It would have also been appropriate to have conducted community analyses throughout the river, making use of kick nets and drift nets in the riffle-runs, emergent insect traps during late Spring and summer months to assess potential prey species (hence exposure) for insectivorous birds and mammals, and potentially gut-content analyses for fish species in the River. It would appear that at least a part of this information does exist, and to the degree practical it would enhance the ERA to summarize this into a qualitative description of the benthic ecosystem. It would also be useful to include in a general discussion any findings related to crayfish, mussels, and dragonflies to the overall description of the benthic communities.

RESPONSE 1-TT-5

Appendix A.1 contains discussion of crayfish (Section III; 2-29 to 2-30), mussels (Section III; 2-7 to 2-14), dragonflies (Section III; 2-14 to 2-20), and other

> invertebrates. EPA will provide additional clarification on the use of this information in the revised ERA. Please refer also to the response to General Issue 4.

Moving onto the actual assessment and measurement endpoints, the stated assessment endpoint was listed as "community structure, survival, growth and reproduction of benthic invertebrates". Benthic invertebrates, by definition, would be animals living on or near the bottom of an aquatic habitat (Thorp and Covich 1991). This is fairly broad and includes organisms living within the substrate (infauna) or upon the substrate (epifauna). A number of those same species will live in or upon aquatic vegetation. It appears that the ERA chose to focus on a narrow subset – principally benthic infauna. This is not only reflected in the overall sampling design, but is also captured in EPA's response to the specific question – define benthic invertebrates (Question TT32 – EPA's Response to Questions dated 12/11/03). EPA's response was that

> "...benthic invertebrates refer to invertebrates that are exposed via sediment and porewater uptake pathways. These organisms include "infauna" and "epifauna" but exclude "water column invertebrates""

This is an important distinction – and should be reflected within the assessment endpoint. Exposure is "via sediment and porewater", and I would submit then by definition the assessment is community structure, survival, growth and reproduction of benthic infauna.

RESPONSE 1-TT-6

> Please refer to the response to General Issue 11.E. EPA concurs with the Thorp and Covich (1991) definition, which is consistent with the definition provided by Reviewers' Questions (December 2003). Although pore water exposure is greater for infauna than for epifauna, exposure of epifauna to pore water PCBs occurs via diffusion from bed sediment. The assessment endpoint did not focus on a narrow subset of benthic invertebrates, but rather included both infauna and epifauna.

In summary, and germane to my subsequent arguments below, (1) the selection of measurement endpoint should be consistent with the route of exposure stated (sediment and porewater), (2) risks to a very important source of aquatic insect prey for fish, birds and insectivorous mammals within the PSA is not addressed by this assessment endpoint, and (3) management actions that could come from this endpoint are only applicable to those areas where soft sediments and benthic infauna exist. In fact, the inconsistency within the assessment and measurement endpoints, along with some specific data quality problems, impair the use of these data in formulating a risk-management decision.

RESPONSE 1-TT-7

> Please refer to the response to General Issue 11.E. The measurement endpoints, which included infauna and epifauna, are consistent with the stated exposure routes, include the aquatic insect prey described by the Reviewer, and do not apply only to soft sediment inhabited by infauna. EPA believes that the Reviewer misunderstands the endpoint definition (see response to General Issue 11.E) and the use of data (see response to General Issue 6.C), and therefore

does not agree that the data are impaired with respect to formulating risk management decisions.

Amphibians

I am generally impressed with the comprehensive set of ecological and toxicological data assembled for this study. I have never seen a system better characterized for frogs in an eco risk assessment anywhere! This is an impressive collection of habitat and population data. The data was presented in the problem formulation in a logic and clear manner. Presentation of a life-history summary with Section 3, and in Appendix E, is something that should be done for all of the assessment endpoints. Studies undertaken by the EPA, and GE, clearly were designed to answer a long-overlooked, and data-poor area of toxicological research.

A consistent theme that I have raised throughout my comments is the very ambiguous formulation of the assessment endpoint, and "survival, growth and reproduction of amphibians was no exception. While not intended as an assessment endpoint, the "effects" box in Figure 4.1-1 was closer to what was probably intended: "Survival, growth, reproduction, and possible community alteration". There was sufficient specificity within the measurement endpoint text box on page E-5 to understand what the intended study was undertaking.

RESPONSE 1-TT-8

Please refer to the response to General Issue 2.

As noted, the ecosystem characterization and Problem Formulation would benefit by clearer links between the two documents. A discussion of the process of selecting amphibians as a Receptor of Concern and the assessment endpoint formulation would greatly aid in the transparency of the overall document.

RESPONSE 1-TT-9

Please refer to the response to General Issue 2.

Fish

For fish the ecosystem has been well characterized and, in general, appropriately applied to the Problem Formulation and subsequent ERA. As with all of the endpoints in this study, there is a tremendous volume of information that is presented in multiple documents. For example, ecological information on the largemouth bass may be found within the Work Plan, the RFI, the species profile for largemouth bass in the accompanying CD, in the Fish Biomass Estimate for the PSA, the GE study, and in Appendix F.2. It would greatly benefit the overall process to append or include the Species Profile – being sure to include all of the information from the various sources within Appendix F. The species profile should strive to include all of the GE habitat information (which it currently does not appear to), and be sure that life history information in Appendix F.2 is consistent with the species profile (see diet information).

RESPONSE 1-TT-10

GE's largemouth bass report was not available when the species profiles were prepared. EPA will give consideration to revisions to Appendix F.2 and the species profiles to ensure consistency among the information used in the ERA to characterize risks to fish.

A figure showing fish collection and tPCB concentrations within the PSA would be helpful.

RESPONSE 1-TT-11

Please refer to the response to General Issue 1.

Insectivorous Birds

For insectivorous birds, the upland ecosystem appears in general to have been adequately characterized, and appropriately applied to the Problem Formulation and subsequent ERA. A deficit in the document is lack of transparency in providing links between the various studies conducted, and that the information collected and evaluated is not adequately represented in the ERA and Problem Formulation section.

Pertinent information is scattered through several documents, and is not drawn together into a single coherent summary – which in this case should be Appendix G. Documents that have to be accessed include the Work Plan (specifically Appendix A.24), Appendix G to the ERA, Final Report for tree swallows by Dr. Custer, the scope and final report on Robin Productivity prepared by ARCADIS for GE, and the landscape and species profiles prepared by Woodlot. There were also several instances where data were collected, and possibly used, but were not found anywhere in the ERA.

RESPONSE 1-TT-12

EPA concurs with this comment and will expand the assessment endpoint appendices to ensure that all the information necessary to understand how the assessment was conducted is appropriately discussed. Also, please refer to the response to General Issue 1.

The rationale for placement of the nesting boxes in the Custer study is not covered in their summary document, but can be found in a careful read of Appendix A.24 in the Field Sampling Plan (Weston 2000). Appropriately, this included adequate tree swallow habitat and accessibility, but a very key consideration not made clear in the ERA reporting is that "the sites cover(ing) a range of sediment PCB contamination in the Lower River between the confluence and Woods Pond". This is a good characteristic, but the selection process is not transparent and should be documented in the final ERA.

RESPONSE 1-TT-13

EPA concurs with this comment. Please refer to Response 1-TT-12.

A corollary to the above is the lack of connectivity between other studies conducted under the ERA to this specific receptor group. For example, it appears that sediment PCB concentrations played a role in determining location of nest boxes, but this is not discussed anywhere in Appendix G or the supporting documents. Appendix A.20 to the Field Study Plan for the ERA (Weston 2000) states that "…WESTON, will collect sediment samples within the tree swallows' expected foraging radius around the next boxes to support this study." If such a study was conducted, it is not discussed in the ERA.

RESPONSE 1-TT-14

EPA concurs with this comment. Sediment samples were collected within the foraging area around the nest boxes, as described in the SIWP. Please refer to Response 1-TT-12.

Tree swallow sightings listed in the Species profile are not discussed in Appendix G.

RESPONSE 1-TT-15

EPA concurs with this comment. Please refer to Response 1-TT-12.

Aquatic insect collections made for the benthic invertebrate studies could/should be discussed in the context of prey items for tree swallows. For example, comparing Figure D.2-1 with Figure G.1-1 shows that benthic invertebrate sampling stations 1 – 3, 6, and 8 all occurred within nest-box locations. Benthic tissue measurements for tPCBs were made at each of these sites, and could/should be discussed within the context of tPCB measures made of prey items collected from juvenile birds.

RESPONSE 1-TT-16

In the case of tree swallows, benthic invertebrate samples collected from sampling stations 1-3, 6, and 8 were not used because of the availability of stomach content samples, which provided a more direct measure of concentrations of COCs in tree swallow food. EPA will clarify why benthic invertebrate samples were not used as part of the exposure analyses for tree swallows and discuss how these data compare to the stomach content samples.

There appears to be additional, important data collected regarding dietary intake for both species of birds evaluated that may not have been included or discussed in the ERA. For example, Slide 12 of Dr. Moore's presentation to the Peer Review Panel on December 18th for insectivorous birds listed dietary composition by dry mass for both tree swallow and American Robin. When questioned about these data, relative to the literature-based tables provided in Appendix G (Tables G.2-1 and G.2-26), Dr. Moore indicated that the data in his slide were site specific. Dr. Custer's independent report also indicates that "70% of dietary items of tree swallows from the Housatonic River were aquatic insects (unpubl. data)". If such data do exist, then the statement in the uncertainty section of Appendix G that states the "dietary compositions were derived from the reported in the literature for birds collected from other geographical locations" needs to be changed. If site specific data was indeed used, then I surmise that the data Dr. Moore used is perhaps that which Dr. Custer refers to in here report. When queried at the Panel meeting, EPA indicated that the unpublished insect data of Christine Custer is not in the ERA, but EPA has it. Recommend that these data be incorporated into the Final ERA.

RESPONSE 1-TT-17

EPA agrees. Site-specific dietary composition data for tree swallow will be added to the revised ERA and the Uncertainty Section of Appendix G will be revised accordingly.

The GE Robin study reports collecting nest locations, but those are not provided in a figure. How the resultant nests compare to soil PCB concentrations, and to tPCBs in the soil invertebrate collections made and shown in Figure G.1-2 should be made. A direct link between exposure and effects (or in this case, lack of effects) should be made.

RESPONSE 1-TT-18

EPA concurs; however, the GE report did not include such a figure. EPA will attempt to develop a figure presenting nest locations and tPCB concentrations in soil and soil invertebrates for the revised ERA, and to provide a discussion of any relationships between the data sets.

This information is available, and the final ERA should draw links from the ERA to the supporting documents. For example;

- Discuss habitat preference and type within the Problem Formulation for both the American Robin and tree swallow. This may be more appropriate for Appendix G. Draw direct links from the Problem Formulation to the supporting information.

RESPONSE 1-TT-19

Text describing habitat preferences of tree swallow and American robin can be found in Section 7.2 and Appendix G.1 and G.2.

- In the text, link selection of nest-box study sites to the tree swallow species profile, and specifically Figure 2 in that profile.

RESPONSE 1-TT-20

EPA agrees with this comment, and will provide a rationale for the selection of nest-box study sites in the tree swallow species profile.

- Provide specific information on how locations for nesting boxes in the tree swallow field studies were selected relative to the species profile (see more below on this).

RESPONSE 1-TT-21

Please refer to Response 1-TT-20.

- Provide a similar figure (Figure 2 – sightings) under the American Robin species profile.

RESPONSE 1-TT-22

EPA concurs and will include a similar figure in the American robin species profile.

- Provide a succinct rationale for selection of study for the American Robin, and provide a figure showing where the sampled nests were located, similar to Figure G.1-1.

RESPONSE 1-TT-23

EPA agrees with this comment, and will provide further rationale for inclusion of the American robin as a representative species. EPA will review the information on nest locations provided by GE and, if possible, will develop the requested figure for inclusion in the revised ERA.

- Provide linkage to other assessment endpoints/studies conducted as part of the ERA. Specifically, sampling aquatic insects (including species and PCB concentrations) and upland terrestrial invertebrates. Ensure that all data collected and used is fully disclosed and discussed in the ERA. (more below).

RESPONSE 1-TT-24

Please refer to the response to General Issue 1.

Piscivorous Birds

Yes, with noting that the same comments raised on the previous four assessment endpoints are applicable here, as well.

Piscivorous Mammals

Yes. It would be appropriate in the Problem Formulation to discuss mink habitat and food preferences, and to show in the PF a figure with mink sightings. All the information exists, but it is scattered through several documents. A single summary would be helpful.

RESPONSE 1-TT-25

Most of this information can be found in Section 9.2 and Appendix I.1 and I.2. Figures showing mink and otter sightings from both the EPA and GE field surveys will be added to the revised ERA.

Omnivorous and Carnivorous Mammals

Same comments from AEs 1 – 4 apply here. No further comment.

Threatened and Endangered Species

General comments given on the other assessment endpoints are applicable here. Would be helpful to have a discussion of eagle sitings and use of the River in the Problem Formulation for this species.

RESPONSE 1-TT-26

EPA will apply all relevant comments from the other wildlife endpoints to the discussion of threatened and endangered species. EPA provided a discussion of eagle sitings and use in the Housatonic River in Section K.2.1.5 of the ERA.

2. COPCS, ASSESSMENT ENDPOINTS, AND STUDY DESIGNS

2. Was the screening of contaminants of potential concern (COPCs), selection of assessment and measurement endpoints, and the study designs for these endpoints appropriate under the evaluation criteria?

Valery Forbes:

Comments: The screening of COPCs was generally appropriate. The use of the pre-ERA to identify COPCs other than PCBs and to determine the downstream boundary beyond which PCBs from the GE facility pose a negligible risk to aquatic biota and wildlife was an effective approach. Nomenclature concerns (Panel Question BS2) could be addressed by referring to the pre-ERA as the Initial Risk Assessment and the ERA as a Refined Risk Assessment. Also, the 3-step tiered approach for establishing an initial COPC list seems to be appropriately conservative with the possible exception of Tier 3 in which evaluation was performed 'subjectively'.

RESPONSE 2-VF-1

EPA believes the terms Pre-ERA and ERA are appropriate as used. EPA agrees that there is some subjectivity associated with part of the Tier III screening. EPA recognizes in many of its guidance documents associated with the ERA process that some level of subjectivity (or professional judgment) is often unavoidable in the decisions made during the ERA development process, but must be clearly described. EPA believes that the thought processes used during the Tier III evaluation were clearly presented and are transparent.

From p. 2-58 the assessment endpoints are defined as representing 'specific ecological values deemed important to protect', whereas measurement endpoints are defined as 'the tools used to determine the outcome for the assessment endpoints'. Although it is possible that some measurement endpoints ma y also be assessment endpoints, in my view the assessment endpoints defined in this ERA (with the exception of community structure) would be more appropriate as measurement endpoints whereas the assessment endpoints would be more appropriately defined as the long-term persistence of populations of benthos, fish, amphibians, birds and mammals in the PSA. To some extent the defined assessment endpoints are redundant. For example, changes in benthic community structure occur because of changes in survival, growth, and/or reproduction of resident species. This is reflected in the WOE for the benthos which states 'the individual measurement endpoints were often applicable to many or all of the assessment endpoints' (D 94) and thus a single WOE was performed that included all benthic assessment endpoints. However, if the assessment endpoints are as stated then benthic toxicity results using different responses (e.g., mortality versus reproduction) should, in principle, have been analysed separately (since they represent separate assessment endpoints) instead of being put into the same analysis. This is probably an issue for other receptors as well.

RESPONSE 2-VF-2

Please refer to the response to General Issue 2.

As is stated by EPA (response to Panel Question JO7), 'Any contaminant-induced response that leads to direct mortality of adult fish, and/or indirect effects on population structure (e.g., loss of recruitment of juveniles to older age classes), and/or health (e.g., reduction in fish growth rates, reduced adult reproduction rates) that lead to an impact on the locally-exposed population [emphasis added] would be considered an ecologically significant response.' This suggests that populations were, in effect, the objects of protection in the present ERA.

I can further point out that populations are specifically named as targets of protection by EPA (1998). When the focus is on the population as a whole, it is acknowledged that a stressor may affect the survival, growth and/or reproduction of some members of the population but that the "acceptability" of the stress is judged in terms of how it effects the population as a whole.

A practical problem with the assessment endpoints as defined is that having several assessment endpoints for each receptor forces the assessor to make judgments as to whether, for example, reproduction, survival, development, maturation, and community condition of amphibians are of equal importance, if the most sensitive of these should drive the risk characterization, or if some should be given more importance than others. An example is for bald eagles where the risk of TEQ was determined to be high for eggs, but low for adults, and the WOE concluded an intermediate risk. Depending on the life-history characteristics of the species, the survival of eggs versus adults may differ in demographic importance. In addition, it is incorrect to assume that high risks for individual performance indicators necessarily and consistently translate into high risks for the population. Clearly EPA recognizes this (see e.g., response to Panel Question MAO2), but have not made the link quantitative. One ecologically based way to weigh risks to different life stages is to consider their importance in terms of population dynamics (e.g., by an elasticity analysis).

For threatened and endangered species the individual is often defined as the protection goal. Partly this is because loss of any or few individuals may have a measurable influence on the population's persistence. However for most other taxa considered, it is persistence of populations, and not individuals, that is the protection goal. Indeed, on page 2-66 it is stated that 'Although many of the endpoints presented are linked to organism-level effects (e.g., survival and reproduction), these endpoints are expected to be strong indicators of potential local population-level effects'. While this is broadly true, the form of the relationships between organism-level effects and population-level effects will vary widely among endpoints and species. Organism-level effects can act as measurement endpoints for estimating population-level effects, but the links should be made quantitative (e.g., through demographic or life-cycle models).

Proposed Changes: I would propose that serious consideration be given to redefining the assessment endpoints: reproduction, growth, and survival as measurement endpoints for the target species considered, and that the assessment endpoints be redefined as 'long-term persistence of populations of receptors'. Likewise it should be clear that for example 'amphibians' are a receptor, whereas Leopard and Wood Frogs are surrogate species chosen to represent amphibians. Also for the other receptors.

RESPONSE 2-VF-3

Please refer to the responses to General Issues 2 and 9.

Thomas W. La Point:

The screening of COPCs was extensive and appropriate. The use of Ingersoll and Long's approaches for estimating threshold effect levels has been shown to result in conservative estimates of sediment-bound contaminant effect levels. I will have more comments, below, on the use of "hazard quotients" in the ERA process. Such quotients are a beginning and should be used only when the variances in denominator and numerator are understood or estimated. Without an estimate of the variances "hidden" by the derived variable, there can be little confidence in assessing degrees of risk in numbers above an HQ of 1. In essence, one could ask "is there any difference in an HQ of 1, relative to an HQ of 5, 7 or even 10?" How can one tell, without knowing the coefficients of variation in numerator and denominator? I would claim that, without this information, one cannot tell.

RESPONSE 2-TL-1

Please refer to the responses to General Issues 7 (interpretation of HQs) and 1.C (probabilistic HQs), and to Response O-VF-4. EPA agrees that the applications of sediment quality values yield conservative indications of potential harm, for both PCBs and other COPCs.

This point was also discussed during the Lenox meeting. There should be greater emphasis on the use of species sensitivity distributions (or exposure distributions), much like has been used for fish (see below). An excellent reference is Posthuma, et alia (2002).

RESPONSE 2-TL-2

Please refer to the response to General Issue 1.C.

To this end, Appendix C.5, concerning point estimates and UCL calculations is excellent and provides an excellent rationale for, and approach to, the use of UCLs in this ERA.

As to the selection of assessment and measurement endpoints, I will have more to say on each one, below. However, one glaring need in this ERA is to include bats. They are present during breeding (Woodlot Report, Attachment C) and are very often prolific insectivores – and often focus on aquatic insects. Bat houses are relatively easy to construct to attract bats (much like the houses for swallows and other avian species) and would have allowed a good estimation of COPC accumulation and effects. I would also liked to have seen more terrestrial invertebrate analysis.

RESPONSE 2-TL-3

An ecological risk assessment for the threatened and endangered small bat, small-footed myotis, was conducted and presented in Appendix K and Section 11. Small-footed myotis is expected to be exposed to high concentrations of tPCBs and TEQ because of its high metabolism. Therefore, it is unlikely that other bat species will have significantly higher exposures. A qualitative assessment was carried out for other bat species by comparing their life history, foraging behavior, habitat preferences, and diet to small-footed myotis. The Indiana bat, northern myotis, and little brown bat, all belonging to the genus

Myotis, were shown to be similar to the small-footed myotis with regard to the factors that influence exposure.

Although in hindsight it might have been preferable to have selected a bat species to study rather than the tree swallow, at the time (1998) tree swallows were thought to be an appropriate receptor, and much of the natural history work on bats had not yet been conducted. Field techniques for studying bats, such as bat houses, were not widely used at that time.

Although sample size was small for terrestrial invertebrates, each sample was composed of many invertebrates, and sampling was conducted across the gradient of contaminant concentrations. Thus, it is possible to use these data to obtain reasonable estimates of central tendency of contaminant concentrations for input into wildlife exposure analyses. EPA agrees that additional terrestrial invertebrate samples would have reduced uncertainty in the exposure analyses conducted for American robin and short-tailed shrew, as would increasing the sample size for virtually any type of data collected for any endpoint; however, funds were not unlimited so choices had to be made. This source of uncertainty is acknowledged in the Sources of Uncertainty sections for insectivorous birds, and omnivorous and carnivorous mammals.

One of the comments brought up during the Public Meeting questioned the extent of sampling downstream from the PSA. It is my understanding that relatively few samples have been collected (n = 66). If this is so (I could not confirm by studying the Corps report (Final Supplemental Information Work Plan) or the Weston July 2003 report), then there should be continued efforts towards understanding the distribution of total PCBs downstream from Reach 9.

RESPONSE 2-TL-4

Several hundred samples were collected downstream of the PSA in Reaches 7, 8, and 9 (43 miles to the MA/CT border). With regard to the extent of sampling in CT, please refer to the response to General Issue 4.

James T. Oris:

Comment 2.1. Screening of COPCs. After hearing and reading answers to questions posed to EPA by the review panel, I believe that the screening process was appropriate. However, the specific sections that describe the screening process and the decisions for what to include in the COC list need to be revised to clarify previous questions from the panel.

RESPONSE 2-JO-1

Please refer to the response to General Issue 3.

Comment 2.2. Selection of assessment and measurement endpoints.

2.2.1 As I have previously stated in my questions, I do not quite understand the wording of the assessment endpoints. Assessment endpoints are supposed to encompass the goals of the risk assessors. Stating an assessment endpoint as "Survival, growth and reproduction of birds" doesn’t' tell me what the desired outcome or goal is to be. The response to my question about this form of statement was that the goal was no impairment of the stated endpoints. A goal of "no impairment" is probably unrealistic given the multitude of potential stressors in a combined industrial, urban, agricultural, and rural watershed. "No impairment due to PCBs" could be a realistic endpoint. "No impairment relative to reference conditions" could be a realistic endpoint. Regardless of what the sidebar-stated goals are of these assessment endpoints, they are not enumerated in the ERA document. The statement "Survival, growth, and reproduction of fish" is a value-neutral statement and is not a well-defined assessment endpoint. This is more a statement of a measurement endpoint to me. In the documents presentation meeting, the EPA's presentation by Sue Svirsky used the following statement for the assessment endpoints (slide 10 of presentation): "Survival, growth and reproductive success of [list of receptors]". Note the word "success" is in this statement and doesn't appear in the ERA document. Additional explanation or revision of the assessment endpoint statements in the document is necessary.

RESPONSE 2-JO-2

Please refer to the response to General Issue 2. The inclusion of the term “reproductive success” on the slide was meant only as a different way of saying reproduction.

2.2.2 It is not clear why some endpoints were more restrictive than others. For example, community structure is a stated endpoint for invertebrates, but not for fish. This is even though fish community structure was analyzed and discussed in the document. In addition, population structure modeling was an endpoint for amphibians, but not for fish. This is even though fish displayed an abnormal age structure and there was evidence of reduced reproduction, and there was apparently little effort devoted to identifying areas of recruitment. In addition, no attempt was shown of field-level assessment of fish health (community or population) with PCB concentration, even though tissue residue data were available and DELT data provided with my questions indicate that there were high levels of deformities, some of which correlated to deformities seen in laboratory studies in a dose-related manner with PCB concentration.

RESPONSE 2-JO-3

Please refer to the responses to General Issues 9 (population endpoints in ERA), 13.E (disease and DELT data), and 13.F (population modeling and demographics of fish). Invertebrates, fish, and amphibians were all considered using measurement endpoints that reflected both individual and population/community level responses. The revised ERA for fish will include additional information/analysis related to field measures of community health. This will include analysis of largemouth bass age structure, field observations of deformities and disease, and population modeling of largemouth bass.

EPA agrees that a concentration-response analysis was not conducted in conjunction with the fish field studies. The study designs did not allow for such assessments, which is reflected in the weighting factors assigned to be field studies. The field studies did allow for a semi-quantitative evaluation of reproduction (e.g. catch-per-unit-effort of YOY bass), and habitat assessments were conducted in areas of potential recruitment. EPA concluded that these studies were relevant for evaluating whether large responses (e.g., reproductive failure) occur in PSA habitat, but were not sufficiently precise to identify either smaller magnitude impacts (e.g., 20% loss of recruitment) or highly localized impacts.

2.2.3 All assessment endpoints need to be re-evaluated in terms of the societal value of the receptors. The decision to use "population-only" level effects (i.e., a reproducing population) may not be appropriate for all receptors. EPA guidance allows for the incorporation of societal value when appropriate. In this risk assessment, this is especially evident with the sections dealing with fish. It is apparent that society places high value on the health of fish. I differ in opinion with the EPA and others on the panel as to what a population level effect is here. To me, a high incidence of disease in a group of fish is a population level effect. As evidenced by public comments -- every single public speaker at the Cranwell meeting said the same thing-- it is clear that the public considers a diseased population of fish (regardless of whether they are reproducing) to be impaired. It is also clear that, even though (as stated by the EPA) the State and Federal agencies that spoke at the meeting participated in the development of assessment endpoint, these same agencies and the public felt very strongly that the fish assessment endpoint was incorrectly defined.

RESPONSE 2-JO-4

Please refer to the responses to General Issues 2 and 9.

Comment 2.3. Study designs appropriate? I do not agree with some of the restrictions that the stated assessment endpoints generated in the design of studies (cf., comment 2.2.2 and 2.2.3), but generally, within the stated assessment endpoints, studies within the PSA were appropriately designed. There are some studies that could have been done differently or some analyses that still may be possible given the designs (e.g., fish deformities v. PCB tissue residue), but overall they appear to be adequate.

Mary Ann Ottinger:

General Comments:

The screening and identification of COPCs and subsequent contaminants of concern (COCs) was thorough and considered a range of possible contaminants. The elimination of some COPCs is rationalized, although not completely convincing in the case of some pesticides, which may occur in a patchy distribution, especially along areas of the river in proximity of agricultural activities. Moreover, relatively small concentrations of some pesticides are known to interact with PCBs and/or with other contaminants, thereby confounding the biological responses observed and potentially obscuring some of the interpretations. For example, lack of consideration of these COPCS at some regions downstream or even in the reference sites (which had low level total PCBs (tPCBs) may result in a masking of the responses, ultimately diminishing the magnitude of difference in a measurement endpoint for the animals in the PSA.

RESPONSE 2-MO-1

Please refer to the response to General Issue 3.

The selection of assessment and measurement endpoints are of concern because most measures are relatively insensitive. Detection of significant differences in these endpoints, primarily mortality and reduced reproduction, with monitoring of survival of young in selected species, will be significant when the indigenous population may be in jeopardy in that region. Biologically relevant measurement endpoints that are responsive to established actions of the constituent PCBs, such as thyroid hormone in some cases, neurochemicals, or behavior would have provided sensitive and reliable indices of exposure and impact that would precede catastrophic impact to a population. The study designs appeared to be limited by the realities of fieldwork. For example, there were only 9 kingfisher nests identified and 3 of these nests were depredated, leaving a very low sample size and questionable reference data. Similarly, the osprey was known to inhabit the PSA, but no field studies were conducted; the ERA rests on models built from information in the literature. This is a reasonable approach, given the lack of field data in the ospreys on the primary study site (PSA), and the models for risk assessment did consider considers data on measured levels of tPCBs in prey species found in the PSA. However, data collected in some additional avian species would have provided a more complete assessment with less uncertainty. Moreover, the field studies in birds would be much stronger if there had been some simply designed laboratory studies to test the impact of the tPCBs known to occur in the PSA. These studies could have been one-generation reproduction tests (refer to OECD avian toxicity test) or egg injection studies in representative species, such as quail or mallard eggs (not chicken as this species is known to be highly sensitive to PCBs) in which known concentrations of tPCBs are administered and then teratological consequences or other selected end points are monitored (see papers by Dave Hoffman for further detail). These types of studies would have clarified the extent of potential consequences due to COCs exposure at selected stages of development and would have verified the conclusions of the field work in the tree swallows within the context of potential sensitivities of ground dwelling birds or other passerines.

RESPONSE 2-MO-2

Please refer to the response to General Issue 2.

Specific Comments [C 6.7-C8.3.25]:

There appear to be some qualitative difficulties with meeting the criteria for the wildlife selected for study possibly due to the availability of representative species. Does this mean that there are insufficient data to meet the requirements of the weight of evidence (WOE) approach? Please rationalize the selection of each wildlife species and ascertain which of the criteria for the ERA have been met to satisfy the WOE approach (see additional specific comments below).

RESPONSE 2-MO-3

EPA provided the rationale for the selection of each of the representative wildlife species in the introduction sections of each wildlife assessment chapter and corresponding appendix. However, EPA will clarify this discussion and also provide a discussion of why other potential representative species were not selected for detailed assessment.

Further, the PSA and the downstream regions appear to be regarded very separately due to the decreasing levels of COCs. This is appropriate because the risk of exposure decreases, but it does not completely eliminate risk, which seems to be the general attitude, especially when considering the issues within the ERA. Therefore, it would be appropriate to have more detailed discussion about appropriate actions, if needed for selected areas downriver (below the ERA) in order to balance potential risk (with runoff or unusual rain events that cause dam opening or other such responses) with the generally low probability of contact with COCs.

RESPONSE 2-MO-4

The concentrations of PCBs in sediment behind dams, which could potentially become mobilized by dam opening or failure, were investigated as described in Sections 4.5.2.5 and 4.5.2.6 in the Housatonic River – Rest of River RCRA Facility Investigation Report prepared by General Electric (BBL and QEA 2003). These data, in conjunction with tissue data and habitat information for the river and floodplain downstream of the PSA (Appendix A.2), were used to formulate MATCs for downstream receptors. The process that was used to characterize downstream risks, based on these data, is described in the ERA. Please refer also to the response to General Issue 5.

Presumably the attribute-scaling factor would be low if the data were collected during one season only (and not during breeding season). Specifically, small mammals were collected in Aug-Sept 1999, mallard and wood duck samples were collected in Aug-Sept 1998, tree swallows were the exception in that they were collected in May-June 1998, 1999, 2000; others including 5 house wrens and 3 chickadees were opportunistically collected in May -June 1999. As a consequence, how was the specific wildlife data scaled and did the lack of multiple years of data for a species result in an underestimation of the attribute scaling? How were the data from the wood ducks used in the assessment?

RESPONSE 2-MO-5

Generally, the scores assigned to an endpoint were higher for the temporal representativeness attribute in the WOE analyses if the field study or samples used in the exposure analyses had been conducted over multiple years and depending upon the appropriateness of the timing of the study. Temporal representativeness, however, is only one of 10 attributes that are considered in weighting a particular line of evidence.

The wood duck samples were analyzed and the resulting data on tPCB and TEQ concentrations used in the exposure analyses for bald eagle and mink. EPA plans to add an assessment of risk to wood duck to the revised ERA, based upon Reviewer comments.

C7.8, line 24-25 what was different about the bird samples?

RESPONSE 2-MO-6

Birds were treated separately from other species due to their high mobility relative to the other organisms. As such, the elevated PCB concentrations in some bird tissues could have been derived from multiple sources including the GE facility.

C.8.3.20 Do any of these sites represent reference sites used in the subsequent studies? How many nest boxes were installed at each site? What happened to the 1998 samples?

RESPONSE 2-MO-7

Reference areas used for the avian surveys and subsequent studies included Threemile Pond Wildlife Management Area and October Mountain State Forest. No nest boxes were installed as part of these survey efforts, nor were any tissue samples collected in 1998 as part of these surveys.

A comprehensive pre-ERA characterization was conducted including an overview of the habitat and species within these habitats. However, in the body of the ERA, little mention is made to this information relative to the selection of representative species for the ERA, except at the end of the report in which there is a general listing of potential effects to other species. The ERA would be strengthened with a discussion of a listing of the species observed in the survey, including estimated population numbers. In the case of field data collected as part of the ERA, information about the sex ratio of the small mammals collected in the traps in both years, with the species density and distribution would also be helpful. These additional considerations will strengthen the logic and support the process used to ultimately identify the representative species and corroborate the conclusions of the ERA.

RESPONSE 2-MO-8

Please see Response 2-MO-3. In addition, Appendix A provides much of the information gathered from the field surveys conducted to support the ERA.

Finally, figures that integrate sediment concentrations, sites of exposure, and species observed at those sites would clearly show the potential relationships of these elements of the ERA.

RESPONSE 2-MO-9

Sediment concentration maps are provided in Volume 1, Section 2.5 and Figures 2.5-5 through 2.5-11; potential sites of exposure as represented by natural community (habitat type) maps are found in Appendix A.1 (Attachments A and J); and species found in each habitat are presented in Appendix A.1, Attachment C. In addition, please refer to the response to General Issue 1.

Brad Sample:

Screening of COPCs

This screening evaluation is among the most extensive and detailed I have seen. This evaluation would pass as a baseline for other sites. An extensive effort was exerted to reduce the list of analytes and focus on primary issues. EPA looked at all data (point by point) using reasonable toxicity values. It should be noted however that this is not a conservative screen.

Soil PRGs for wildlife from Efroymson et al. are based on LOAELs and are not really suitable as screening values. No real change needed - just needs to be noted that screening that results from these values is not conservative.

RESPONSE 2-BS-1

EPA will provide additional text in the revised ERA describing the level of conservatism associated with the use of LOAEL-based PRGs.

Background screening was performed for organics. This is generally not performed. Organics are almost exclusively anthropogenic, so there should not be any present in background. Background screens are generally only conducted for inorganics that, because they are derived from underlying geology, may occur naturally. Suggest adding brief text explaining that the screening of organics is non-standard and discussing why it was conducted.

RESPONSE 2-BS-2

EPA will provide additional text in the revised ERA explaining the screening of organics in the Pre-ERA.

Appendix B, Page B-22, Section B.4.3: What is the basis for the 10% criteria for exceedences under the 'unlikely' category?

RESPONSE 2-BS-3

The 10% criterion used for benchmark exceedences was established based on best professional judgment. This will be clarified in the revised ERA.

Assessment and Measurement Endpoints

As currently presented, the assessment endpoints are poorly defined. This issue carries through the whole report and affects how data are used and interpreted. Based on EPA guidance (EPA 1998), a complete definition of assessment endpoints has two components, 1) entities to be protected and 2) attributes that may be affected or at risk. Implicit in this definition is some level of ecological organization (i.e., individual-, population-, community-, or ecosystem-level) for which protection of the assessment endpoint is sought. None of these components are clearly defined for any of the assessment endpoints described in Section 2.8 nor in Table 2.8-1. For individual entities (identified as receptors in Table 2.8-1; e.g., benthic invertebrates, amphibians, fish, etc.), the attributes (identified as assessment endpoints in Table 2.8-1) more than one level of ecological organization are implied.

For example, the 'assessment endpoints' for benthic invertebrates are community structure, survival, growth, and reproduction. Whereas community structure is a community- or ecosystem-level entity, survival, growth, and reproduction are individual- or population-level entities. Different data (measurement endpoints), different analyses, and potentially different conclusions may be drawn depending on the level of ecological organization evaluated. I would recommend that the level of ecological organization considered for each assessment endpoint be explicitly stated to ensure that it is correctly matched with appropriate measurement endpoints and so that suitable conclusions can be drawn.

RESPONSE 2-BS-4

Please refer to the response to General Issue 2.

Note that the Sediment Quality Triad is not really a measurement endpoint in and of itself, rather it is a risk characterization approach. Each of the components of the triad are measurement endpoints though. Suggest removing it from table and discuss in text only.

RESPONSE 2-BS-5

In its most narrow definition, the Sediment Quality Triad (SQT) consists of measures of chemistry, toxicity, and indigenous biota, with a weight-of-evidence (WOE) approach used to interpret the data. In the Housatonic River application, a broader definition of the SQT was applied, which included other measurement endpoints used to strengthen the cause-effect linkages in the WOE. These additional endpoints included assessments of biomagnification (i.e., bioavailability measures), toxicity identification evaluation (TIE), and assessment of sediment stability, etc. (Grapentine et al. 2002; Burton et al. 2002). Rather than identify all the individual measures contained within the SQT, and to ensure consistency with the SIWP, the broader Sediment Quality Triad concept was used in the summary tables. To address the Reviewer's comment, a footnote will be added to the table to indicate the additional tools used in the SQT applied to the Housatonic River site.

Swallows and robins and shrews and foxes do not represent the same ecological exposures. Whereas swallows are aerial insectivores, robins are ground insectivores. Similarly, whereas shrews are ground insectivores with small home ranges, foxes are wide-ranging predators consuming a diverse diet. Pooling these receptors together oversimplifies and misrepresents the exposure pathways being evaluated. I would recommend that these, and the other wildlife species (piscivorous birds and mammals, special-status species), each be evaluated as separate assessment endpoints. Conclusions about potential risks to different ecological guilds can then be made following the weight-of-evidence for each individual species. This is more appropriate given that the level of assessment is at the local population level, not the community level.

RESPONSE 2-BS-6

Please refer to the response to General Issues 14.C, 15.A, and 16.A.

Volume 1 page 2-66 lines 12-17: This paragraph is of great importance. It is the first time that the level of ecological organization for which the assessment is being conducted is discussed. It also is the first (only?) time at which the disjunction between the level of organization associated with the available data and the level of organization for the assessment endpoint is discussed. This text needs to be greatly expanded and integrated (referenced?) into the discussion for each assessment endpoint.

RESPONSE 2-BS-7

Please refer to the responses to General Issues 2 and 9.

Ralph G. Stahl:

I believe the process of screening the COPCs was appropriate; however, I am not convinced that PAHs and perhaps mercury or other contaminants were not partially responsible for some of the observed effects in aquatic receptors.

RESPONSE 2-RS-1

Please refer to the response to General Issue 3.

I do not believe that the screening approach in the ERA requires substantial revision, but I recommend that the authors temper their statements with respect to tPCBs and TEQ being the causative agents of malformations observed in fish and amphibians. In this regard the screening of COPCs is not consistent and potentially not objective. It is reasonable however given the magnitude of risk question at hand where all the possible COPCs simply cannot be evaluated effectively.

RESPONSE 2-RS-2

Please refer to the response to General Issue 3.

I believe Dr. Oris has pointed this out in his semi-quantitative analysis of the HQs resulting from stressors other than tPCBs and TEQ. In some cases COPCs produced HQs greater than 1, yet these were not discussed in any systematic manner in the ERA, nor was the potential for these other COPCs to produce the morphological abnormalities noted in some of the receptors.

RESPONSE 2-RS-3

Please refer to the response to General Issue 3.

I believe the assessment and measurement endpoints were, in general, selected properly. As others on the panel have noted, there is some concern on the wording of the assessment endpoints. For example, an assessment endpoint should not be so overly broad that one is incapable of linking the measurement endpoint results to potential risks. Most, if not all, of the 8 assessment endpoints are written very broadly and this becomes problematic in the final assignment of risk in the WOE process. I recommend that the authors reconsider the wording of the assessment endpoints to reflect more specifically those attributes of the assessment endpoint (or receptor) that are most valued and worthy of protection. In this way the final assignment of risk in the WOE process will be more evident given the measurement endpoint results presented in the ERA.

RESPONSE 2-RS-4

Please refer to the response to General Issue 2.

Based on some of the EPA presentations, it appears that some information was collected on receptors (or assessment endpoints) yet these did not warrant further evaluation in the ERA. Two examples are the dragonfly and waterfowl. It is not clear from the ERA why these two receptors, or assessment endpoints were not included even though preliminary data were presented on them. Even more puzzling is the fact that a waterfowl consumption advisory is listed for some portions of the Housatonic River due to the potential for tPCB contamination in edible tissue, yet waterfowl are not given any substantive discussion or evaluation in the ERA. If this point has been covered in the human health risk assessment for the Housatonic River, it should be so noted in the Housatonic River ERA. If not, I recommend that the authors of the ERA more fully explain the reasoning for not including such receptors as the dragonfly and waterfowl. In this regard, the selection of receptors does not appear to be consistent.

RESPONSE 2-RS-5

Please refer to the response to General Issue 4.

Some of the measurement endpoints resulted in data sets that were limited. For example, there were limited numbers of fledgling piscivorous birds (belted kingfisher) observed in the field study conducted by GE, yet this was the only field investigation available for use in the ERA. The bulk of the ERA relative to the piscivorous birds was based entirely on modeling, and therefore carries a substantial level of uncertainty.

RESPONSE 2-RS-6

EPA concurs with this comment. The uncertainty associated with the assessment of risk to piscivorous birds is acknowledged in the Sources of Uncertainty sections in Chapter 8 and Appendix H.

In another example, few mink and otter were observed in the field despite significant effort both by EPA and GE, resulting, again, in sparse data sets. The lack of high quality field data became the underlying reason for the uncertainties and conflicts noted in the conclusions on potential risks.

RESPONSE 2-RS-7

Please refer to the response to General Issue 8.

With this in mind, it is important given the circumstances and the significant efforts undertaken by EPA and GE, that all data collected in the PSA and elsewhere be utilized to the maximum extent possible. Some data sets will be sparse owing simply to the difficulty in obtaining the organisms that are the subject of the study. These should, nevertheless, be utilized as is appropriate in helping to understand the potential for risks to ecological receptors in the Housatonic River and its watershed. Simply dismissing the studies out of hand, or on the basis of the absence of prior review of the design is not prudent given the magnitude of the risk question at hand. Therefore it is reasonable to include as much of the data as is possible, and identify those strengths and weaknesses of those data in the ERA.

RESPONSE 2-RS-8

EPA does not believe that any studies were dismissed out of hand or because of absence of prior review of study design. In a few cases, analyses or new data were generated on behalf of GE following release of the current ERA for review. As a result, they could not be included in the ERA. Such analyses and data will be considered for inclusion in the revised ERA.

Timothy Thompson:

Benthic Invertebrates

The use of sediment quality thresholds and water quality criteria for screening COPCs are very well established in the scientific and ecological risk assessment community. Methods used by the ERA for screening are appropriate.

As presented above, the ERA's intent was to assess risks to infaunal organisms – not to aquatic invertebrates in general. That the ERA's intended is to assess benthic infauna is reflected within the Supplemental Investigation Work Plan (Weston 2000). Section 7.2.4 poses the question "are ambient PCB levels in sediment sufficient to cause adverse effects in benthic organisms..". Within Appendix A.13, it states that "Sampling of benthic macroinvertebrates inhabiting soft (depositional) sediments will be conducted at 13 locations in the Housatonic River watershed".

RESPONSE 2-TT-1

> Please refer to the responses to General Issues 11.E and 11.F, and to Response 1-TT-7. EPA does not agree that the intent of the ERA was to assess infauna only. The 13 stations were selected to provide a representation of the broad habitat types available with the PSA reaches; therefore, cobble/riffle habitats, which are rare in the PSA, were not included.

What is very clear and important is that the ERA never intended to assess overall benthic infaunal community health – rather it set out to see if a dose/response relationship could be established between levels of PCBs and the specific measurement endpoints. Section 2 of Appendix A.13 to the Work Plan states that

"The benthic macroinvertebrate investigation is sharply focused on providing information specific to the three data quality objectives outlined above, rather than providing a full characterization of the benthic macroinvertebrate community, which is beyond the scope of this study." (emphasis added)

Those objectives were defined as evaluating PCBs, other contaminants and effluent from the WWTP, providing tissue concentrations of benthic macroinvertebrate groups for use in the modeling, provide "one of the three components of the Sediment Quality Triad approach to evaluating the potential ecological risk across a range of PCB concentrations in the sediment."

RESPONSE 2-TT-2

> EPA's intent was to assess overall benthic community health, to the extent practicable. However, the large size and complexity of the PSA introduced practical considerations for the evaluation of benthic invertebrates; therefore, the statement in the Work Plan was added to make this clear. A representative subset of habitats within the PSA was chosen for the application of the sediment quality triad, and concentration-response relationships from these locations were used to extrapolate risk estimates to other PSA locations. The quote from Appendix A.13 of the SIWP was meant to convey that all subreaches and habitat types could not be practically evaluated. An example of the above is that cobble/riffle habitats were not sampled in the EPA program (but were the focus of a historical GE study). As a result, organisms preferentially associated with these habitats (e.g., stoneflies) were not directly assessed in the EPA study.
>
> Because EPA focused on sample locations that are representative of the broad substrate and habitat conditions of the PSA, the findings of the sediment triad evaluation are generally applicable to other areas of the PSA. Extrapolations of concentration-response relationships to microhabitat types that were not directly assessed by EPA for toxicity and community data carry larger uncertainty. However, these areas comprise a small percentage of the PSA sediment bed,

and play a relatively minor role in the overall ecological function of the PSA. This will be discussed in the uncertainty assessment in the revised ERA.

Further support that measurement endpoints were selected not to assess overall community health is found in Section 3.2 of Appendix A.13 which states that

"The rationale for selection of the 13 locations to be sampled in the benthic macroinvertebrate study is presented in Subsection 2.1.2. The locations are not intended to be representative of the entire river but rather are intended to encompass the range of sediment PCB concentrations in the Lower River between the Confluence and Woods Pond."

> "Of the remaining five target stations, four will be selected to (1) provide data from locations with PCB concentrations that expand upon the range of concentrations represented at the Triad locations and (2) investigate the potential effects of the WWTP discharge on benthic communities in the river."

This argument is crucial to supporting the assessment – if a dose/response relationship cannot be demonstrated within the measurement endpoints, then there is very little utility in the acquired data. Based upon my read of this information, the assessment endpoint would be reformulated into the following testable hypothesis: There is a dose-dependent response to PCBs in community structure, survival, growth and reproduction of benthic infauna. That is the way the study is currently constructed, so this would be a more appropriate assessment endpoint.

RESPONSE 2-TT-3

Please refer to the response to General Issue 2 (Assessment Endpoint development) and Response 2-TT-2. EPA agrees that establishment of concentration-response relationships was an important objective of the ERA. EPA also believes that both the toxicity data and benthic community data support the observation of concentration-response relationships, even with the issue of small scale variability.

The target sampling locations were intended to be representative of the entire PSA. The locations were selected both to bound the range of PCB concentrations observed and to provide representation of the softer sediment in the dominant habitat types. Selection of representative locations and the selection of PCB concentration ranges are not mutually exclusive objectives. It is for this reason that the 13 locations included representation of both the relatively coarser-grained habitat typical of the upstream portion of the PSA, and finer-grained habitats found in the downstream portion of the PSA, and included samples across the entire geographical span of the PSA. Due to the high variability in sediment PCB concentrations, the measured PCB concentrations in fine-grained sediment during the benthic community assessment were somewhat lower than expected, based upon the sampling conducted to identify appropriate stations across the geographical and concentration range.

Having noted that, I am in agreement with some of my Panel colleagues who noted that the Assessment Endpoint should be the long-term persistence of benthic invertebrates. Dr. Forbes' comment that "…consideration be given to redefining the assessment endpoints: reproduction, growth, and survival as measurement endpoints for the target species considered and that the assessment endpoints be redefined as impairments to long-term persistence of populations of target taxa". Taking this approach would, however, require more of a reliance on community structure than the bioassay results.

RESPONSE 2-TT-4

Please refer to the response to General Issue 2.

I take the dissenting position from EPA and some of my Panel colleagues in that I do not agree that measurement endpoints centered on benthic exposure represent a "worst case exposure", which would be representative of all infaunal and epifaunal species. The risk assessment is about important receptor groups and the appropriate receptor pathways. To state that results of chironomids bioassays are indicative of protection to EPT or odonates epifaunal species, or to state that daphnid bioassays are applicable to benthic infauna, does not provide a direct link between exposure, effect, and importance to the food web on the Housatonic River. On that note, the Panel was generally in agreement that daphnids were not appropriate measurement endpoints to the stated assessment endpoint.

RESPONSE 2-TT-5

The term "worst case exposure" implies the exposure pathway that causes the largest accumulation of contaminants at the site of toxic action. The field data and bioaccumulation model predictions indicate that invertebrates associated with the sediment bed have the highest exposures to PCBs and other COCs. Furthermore, the broad definition of "benthic invertebrates" applied in the ERA encompasses a wide range of feeding strategies and combinations of exposure pathways. Chironomid larvae reflect exposures dominated by the sediment bed, *Hyalella* reflect a mixed water-column/sediment exposure, and *Daphnia* reflect a mixed water-column/sediment exposure with increased contribution of the water column. EPA believes that all three organisms are representative of the Housatonic River communities, including odonates and Ephemeroptera, Plecoptera, and Trichoptera (EPT) taxa. The Reviewer's comments regarding receptor pathways are based on incorrect assumptions regarding the definitions applied in the ERA (see response to General Issue 11.E).

Please refer to the response to General Issue 11.E regarding use of daphnids as valid measurement endpoints in sediment toxicity testing.

A final note, it is interesting to note that the list of assessment and measurement endpoints between the Work Plan (Section 7) and the ERA are different. Table 2.8-1 adds comparison of tissue chemistry to tissue effects thresholds, and drops survival and physiology of freshwater mussels. An explanation for this should be included in the final ERA. Tissue residue values will be discussed more below.

RESPONSE 2-TT-6

Please refer to the response to General Issue 2.

Amphibians

Given the general lack of amphibian tissue toxicity reference values, use of SQVs and AWQC are not an inappropriate way to screen for COPCs. However, since tissue measurements of Appendix 9 metals, PAHs, organo-chlorine pesticides were apparently made for field-collected frogs (based on a submitted chain-of-custody report), better use of these data should have been made. In general, the results shown in Appendix C to the Rana pipiens study showed non-detects for most organics, but DDE was detected, but not discussed.

RESPONSE 2-TT-7

EPA will provide a more detailed justification for the COCs incorporated in the amphibian Assessment Endpoint, including a discussion of the tissue concentrations.

Fish

Screening of COPCs in the pre-ERA are appropriate under the evaluation criteria and are consistent with existing scientific knowledge and practice for ERAs.

What was not appropriately applied was pointed out under my general comments at the beginning of my review; assessment endpoints are ambiguously worded, and are not specific, testable hypotheses. As listed in EA Table 2.8.1, the assessment endpoint for fish is "survival, growth and reproduction". In this form, the assessment endpoint allows for labeling any effect observed on these endpoints as being "adverse". The risk question posed in Section 7 of the Work Plan offered somewhat more specificity;

> Fish Community—Are ambient levels of PCBs present at concentrations in fish habitat sufficient to cause adverse effects in fish in the Housatonic River?

Even in that form, there is no means of supporting an "appropriate managerial decision". As will be discussed below, the ERA does not make the distinction between potential individual effects as indicated in the laboratory toxicity studies, and a determination that there are ecologically significant effects to populations of fish in the PSA.

RESPONSE 2-TT-8

Please refer to the responses to General Issues 2 and 9.

As is aptly done in the bird and wildlife sections, a statement of probability of exposure is coupled with a probability of effect. For example, for risk management purposes a statement such as "there is a 20% encountering PCB concentrations at which 50% of the exposed fry will experience mortality", coupled with an understanding of field-measured populations, allows for setting what will be appropriate management actions for this endpoint. While probability of exposure is indeed presented, the MATC used is a mixture of lethal and subchronic effects in a fashion that cannot be used to integrate the data with population-level measurements.

RESPONSE 2-TT-9

Please refer to the responses to General Issues 7 and 1.C.

There was considerable discussion among the Panel about the appropriateness of using the presence of disease and deformities as indicators of population health. On one level, the incidence of disease and deformities does not necessarily equate with population decline, but the presence does indicate presence of a stressor. I believe these data may be appropriately used, but either the assessment endpoints need to be broadened to include a "no mutated, diseased or ugly fish" rule on the River, or the measurement endpoint of reduced survival must be tightly linked to disease and deformities.

RESPONSE 2-TT-10

Please refer to the response to General Issue 2.

Insectivorous Birds

For insectivorous birds, the COPC screen, selection of assessment and measurement endpoints, and the resultant study design for these endpoints followed and objective and scientific process. The study plans and designs are in general within accepted scientific practice. Specific issues for insectivorous birds are discussed under Charge Question 3(a), below.

Piscivorous Birds

Screening of COPCs were appropriate for this receptor.

The same general comment concerning transparency of selection of this assessment endpoint – and the accompanying measurement endpoint, with the resource trustees is applicable here.

RESPONSE 2-TT-11

Please refer to the response to General Issue 2.

EPA's study design using principally modeled exposures is appropriate and has been previously commented on. The documented presence of osprey within the PSA demonstrates the appropriateness of utilizing this species as a receptor of concern.

As has been noted previously, the GE field studies were not part of the a priori determination of measurement endpoints, and here to the principal aim was solely to document presence of reproducing kingfishers – and is not necessarily a demonstration that reproductive impairment is not occurring.

RESPONSE 2-TT-12

EPA concurs with this statement.

Piscivorous Mammals

Yes – this is a good study with solid consideration of the study design and intended use

Omnivorous and Carnivorous Mammals

Same comments from AEs 1 – 4 apply here. No further comment.

Threatened and Endangered Species

Yes.

3. EVALUATION OF ASSESSMENT ENDPOINTS

3. For each of the 8 assessment endpoints evaluated in the ERA (listed in Attachment B, and for which a specific Section and Appendix was prepared), address the following questions:

James T. Oris:

NOTE: Because of the large amount of redundancy between chapters and appendices, and because more detail is presented in appendices, my comments will reference pertinent sections in the appendices only. Appropriate linkages will need to be made between appendix and chapter.

Mary Ann Ottinger:

General Comments:

The process for the ERA included identification of representative species, conduct of exposure assessments, effects assessments, and risk characterization. This sequence is clearly laid out in the Executive Summary; however, it is sometimes difficult to follow the flow in the body of the ERA. For example, the field surveys are intermixed with the laboratory studies in the amphibian or benthic invertebrate community sections. As such, it becomes difficult to follow the logical sequence, through to the WOE. The presentation made by the EPA, including the GE data, served to integrate the flow of these field and lab studies much more logically and was more convincing in reaching a reasonable WOE conclusion. The flow charts used are excellent and clarify the approach (ex. I.5-7; Figure 5.1-4).

RESPONSE 3-MO-1

EPA agrees and will clarify the logical flow of the assessment leading to the weight-of-evidence for all endpoints, but in particular for the benthic invertebrates and the amphibian endpoints.

As discussed below, the survey demonstrate that amphibians provide an excellent potential index of exposure by considering species density and populations; unfortunately there is a lack of detailed data from more than one species of frog.

RESPONSE 3-MO-2

Wood frogs and leopard frogs were chosen as representative species in the ERA because of their presence in the PSA in different habitats with varying concentrations of PCBs. As noted by the Reviewer, the leopard frog study did not provide the level of detailed data that was obtained in the wood frog study.

However, EPA believes that there is a large body of information that informs the determination of risk to amphibians. In addition to the site-specific studies, toxicity data from other studies in the literature were available for comparison. As discussed in Appendices A.1 and E, a number of different species of amphibians occur in the Housatonic River and floodplain. As described in Section 12.4 of the ERA, risks to other species not explicitly considered in the risk assessment were developed by considering four major factors: diet; foraging behavior and home range; size, metabolism, and life history characteristics; and sensitivity to COCs. These factors, in conjunction with other detailed site-specific data including the amphibian field studies, were used to develop risk estimates for amphibians using the Housatonic River and floodplain.

The studies on fish are the most complete and include both field and laboratory studies. In addition, the sentinel species are representative of bottom feeders, and water column dwellers, selected to include varied habitats within the ecosystem. This is appropriate due to the primary contamination of the aquatic environment; thereby making aquatic species potentially subjected to higher and more sustained exposure, especially the benthic bottom feeders. Predatory species also would be expected to receive higher exposure due to biomagnification and bioaccumulation.

Studies in insectivorous birds should have included other representative species. For example, samples, such as turkey or duck eggs were opportunistically collected, but the results were not utilized in the ERA, even from the standpoint of some additional information.

RESPONSE 3-MO-3

EPA plans to add evaluation of the wood duck to the revised ERA.

In wildlife, as in other categories, assessment endpoints are limited by field constraints. Consideration of impact on wildlife is primarily determined in field studies (with the exception of the mink study), or construction of models based on available literature. It would have been valuable to have additional studies on captive wild species taken from the PSA or a mallard duck study conducted using materials from the PSA, similar to the mink study. In the case of the osprey, the model is entirely based on the literature.

RESPONSE 3-MO-4

Field studies on captive wildlife species are indeed valuable and can significantly improve understanding of risks while reducing uncertainty. However, such studies are difficult to undertake and expensive. In the ERA, field studies were included for benthic invertebrates, leopard and wood frogs, fish, tree swallows, kingfishers, American robin, mink, and short-tailed shrews. Thus, field studies were conducted for most assessment endpoints. Field surveys were conducted for an even broader range of organisms. A limited study of Wood ducks is being added to the revised ERA. Given the practical limitations, EPA does not believe it is necessary to conduct additional field studies to support the ERA.

The ERA evaluations for osprey, red fox, and the threatened and endangered species were based on one line of evidence: modeling of exposure and effects. In these cases, however, the data used to estimate prey concentrations were from samples collected in the PSA and downstream of Woods Pond. Additional lines of evidence would have been difficult to collect for these species, given the inherent difficulties in performing captive studies on these wide-ranging and/or rare species.

3.1 Benthic Invertebrates

3.1.(a) Were the EPA studies and analyses performed (e.g., field studies, site-specific toxicity studies, comparison of exposure and effects) appropriate under the evaluation criteria, and based on accepted scientific practices?

Valery Forbes:

The sediment quality triad approach is a potentially powerful one for assessing risks to benthic communities. Environment Canada has developed a very useful guide to interpreting results of triad assessments, particularly when the different lines of evidence give conflicting conclusions (Reynoldson et al. 2002, HERA 8:1569-1584). There are also other relevant papers in this special HERA issue (2002, volume 8, no. 7) on WOE in sediment risk assessment.

RESPONSE 3.1-VF-1

The ERA authors are familiar with the literature cited by the Reviewer, and some authors contributed directly to the special Human and Ecological Risk Assessment (HERA) issue (Chapman et al. 2002). The ERA followed principles described in the special HERA issue.

Thomas W. La Point:

By-and-large, yes.

James T. Oris:

In general, studies and analyses were done as described and meet accepted scientific practice.

Mary Ann Ottinger:

Overall, the EPA studies were appropriate, with the following comments. Sediment grain size and habitat sediment characteristics contain inherent bias, which may affect conclusions. WOE concluded intermediate to high risk, which decreases with distance from source; this conclusion depends on no further movement of sediment into the flow of the river with exposure of downstream populations. The variables in the system, including sediment grain size and resultant variation in the distribution of contaminants of concern (COCs) increase the uncertainties in the interpretation of the data and make assessments less reliable.

RESPONSE 3.1-MO-1

Data on physical parameters such as grain size or other habitat characteristics would contain inherent bias only if there were a bias in the field or laboratory analytical methods. Field studies were conducted following standard practices and were documented in the work plan and SOPs. There is no evidence of bias in the analytical methods for particle size distributions, sediment organic carbon, or other PSA habitat parameters. The changes that were observed in the substrate parameters, moving upstream to downstream, are reflective of changes in the hydrological and hydrodynamic regimes (particularly the influence of Woods Pond Dam), not sampling bias. The sampling conducted by EPA in conjunction with the triad studies generated sediment habitat information that is consistent with the characterization of these reaches provided in the RFI (BBL and QEA 2003). Although the changes in substrate introduce uncertainty in the assessment of benthic communities, these changes are representative of the conditions in the river, and are not indicative of bias.

The conditions observed in the river system during the investigations represent a dynamic equilibrium that has been reached after a period of nearly 70 years since the beginning of the discharge of PCBs to the system. Risk conclusions drawn from the WOE do not assume a lack of movement of sediment into the flow of the river; it has been occurring and will continue to occur. Extrapolations of risk to downstream areas were performed using the MATCs generated from the data collected in the PSA. As noted for conditions in the PSA, the conditions in the river below Woods Pond have also reached a state of dynamic equilibrium controlled to a large extent by the presence of dams and the topography of the watershed. These conditions are also expected to remain relatively stable unless there is an action, such as dam removal, that occurs. In such a case, the potential for release of sediment and contaminants, and the impact of such a release, will be evaluated.

EPA agrees that the natural variation observed in measures of exposure introduces some uncertainty in the development of concentration-response relationships. However, in the ERA these uncertainties are not so large as to render the assessments unreliable. The uncertainty in concentration-response relationships for individual measurement endpoints is offset by the general concurrence in findings across numerous endpoints. This degree of uncertainty is commonly encountered in risk assessments, requiring that multiple lines of evidence be considered. It is for this reason that the sediment quality triad and WOE approaches were adopted for the ERA. Please refer also to the responses to General Issues 11.B, 11.D, and 11.G.

Brad Sample:

My background in the evaluation of potential effects on benthics invertebrates is limited. However, the studies, methods, and approaches used by EPA appear to be suitable and appropriate.

Ralph G. Stahl:

The EPA field studies were appropriate under the evaluation criteria and based on accepted scientific practices. The field studies, while limited, were sufficient to estimate potential impacts on benthic invertebrate abundance and diversity. The field studies suggest a limited depression, if any, of abundance and diversity, and there appears to be no significant concentration-response between tPCBs and abundance.

RESPONSE 3.1-RS-1

EPA believes that the information presented in the ERA demonstrates that abundance and diversity are significantly related to PCB concentrations. In coarse-grained sediment, statistically significant differences between contaminated and reference stations were observed; these were presented in multiple statistical analyses (MDS, rank plots, ANOVAs, etc) that indicated statistically significant differences related to PCB concentrations. In terms of effect size, the taxonomic richness at coarse-grained contaminated stations is, on average, less than half that of reference stations. The reduction of invertebrate abundance is even more pronounced (approximately one-third that of reference stations). In fine-grained sediment, the relationship is weaker. However, a significant relationship between PCB concentration and benthic diversity measures was observed. This is true of both the Shannon-Wiener Index presented in the response to the Reviewers' questions, and the Simpson's Index evaluation that will be included in the revised ERA. Statistically significant concentration-response relationships were also observed for specific indicator taxa (e.g., chironomids).

In the GE presentation to the Peer Review Panel at the public meeting (Page 14), data from reference stations were not included in the graphics presented, thereby inappropriately eliminating the low concentration end of the concentration-response analysis. This is of importance for the analysis because the observed biological responses are not linearly related to PCB, but rather demonstrate a threshold response. Given that the among-station variability in PCB concentrations in contaminated coarse-grained locations is small relative to the within-station variability (ERA Figure D.2-10), it is not surprising that the statistical relationship of abundance to PCB concentration is weak when the reference data are truncated in this manner.

To address the Reviewer comment, EPA will expand the analysis of concentration-response for benthic community endpoints in the revised ERA, following the recommendations of the Reviewers and incorporating additional analyses that were provided to Reviewers by GE and EPA in December 2003. Please refer also to the responses to General Issues 11.A, 11.B, 11.D, and 11.G.

The site specific toxicity studies in and of themselves were appropriate under the evaluation criteria. However, on further review, the most synoptic exposure values for tPCBs in the sediment toxicity studies appears to not have been used and therefore fail to meet the criteria of appropriate and objective. In my recommendations I have noted the need to revise the evaluation so that the most synoptic data are utilized.

RESPONSE 3.1-RS-2

Please refer to the response to General Issue 6.C. EPA did perform the analyses using the most synoptic exposure values (see Attachment D.5 of the ERA).

Comparison of effects in the site-specific toxicity studies appears to stray from accepted scientific practices. In some cases a comparison was made between the laboratory control and the treatments whereas it is standard practice to make this comparison between the reference area (reference control) and the treatment. Where the authors choose to make comparisons using both the reference area and the laboratory control, or other "control" this should be described clearly in the ERA. Where only one approach was used, then this too should be clearly stated in the ERA so there is no ambiguity in or mis-interpretation of the analysis.

RESPONSE 3.1-RS-3

EPA agrees with the Reviewer that comparisons between the reference areas and treatments should form the basis for MATC derivations and concentration-response analyses. This approach was used in the ERA. For example, page D-50 of the ERA states:

> "Interpretation of statistical comparisons to negative laboratory controls may have some limitations because of the differences in test performance between negative controls and reference sediment. There were several instances for which responses were observed in sediment at one or both of the reference locations, resulting in calculation of very low effects thresholds based on comparison to negative controls. Because differences in responses between negative control and reference sediment cannot be explained on the basis of PCB concentrations … it is appropriate to make comparisons between contaminated stations and reference locations."

The approach used for comparison of effects in the site-specific toxicity tests was consistent with accepted methodologies, and was applied consistently throughout the document. EPA calculated endpoints based on comparisons to the negative control and the two reference sediment samples, to provide the most complete data set, but then used only the reference sediment comparisons for development of the MATC. In the risk characterization, indications of potential for harm "were standardized to appropriate background conditions (e.g., toxicity endpoints were compared to reference Stations A1 and A3 rather than negative control sediment)."

EPA agrees that inclusion of the EC20 and EC50 values based on negative control comparisons in the summary text boxes may have created the impression that these values were used in MATC derivation and risk characterization. This will be clarified in the revised ERA to remove the potential for misinterpretation; however, based on EPA guidance (EPA, 2000a) and toxicity test protocols (EPA/USACE, 1998; PSDDA, 2000; Environment Canada, 1998), EPA believes that a presentation and discussion of negative control performance in the effects assessment is a valid and appropriate component of a comprehensive ERA and will retain this presentation in the revised ERA.

Timothy Thompson:

Within the context of defining risks to PCBs to benthic infauna based upon a dose/response relationship in the legs of the Sediment Quality Triad, the planned field studies, toxicity studies, and accompanying measurements were are all accepted scientific practices. There are elements within each of the "legs" that influence the quality and use of the findings, are discussed below.

Sediment Measurements of tPCBs

The variability and range of tPCB levels in sediment are perhaps the most greatest impediment to the utility of these data for the assessment and characterization of risk. The high spatial and temporal variability in sediment PCB concentrations created difficulties for the ERA writers in that the number of the chemical measurements could not be easily matched with toxicity and/or community structure information. I agree with Dr. Forbes that the "difference in sediment concentration trends (stations 4 – 8) between the benthic community samples (sediment PCB concentration declines) and the toxicity station samples (sediment PCB concentration increases) is unfortunate and does not increase the clarity of interpretation."

RESPONSE 3.1-TT-1

Please refer to the responses to General Issues 11.B and 11G for discussion of issues related to variability in PCB concentrations and linkage to concentration-response.

This is clearly evident by the amount of time that the ERA authors spend trying to grapple with that specific issue – and in particular using the "most synoptic" data point from lab studies vs. the use of an average concentration of the field measurements. This appears to have been a post priori consideration – the Work Plan Appendix A.13 clearly defines analytical sampling procedures and the expectation that the sub-cores collected at each of the sites would be used for characterizing the analytical portion for all the legs of the triad. The "most synoptic" discussion is very difficult to understand as constructed in the document. What is very clear, and very important, is that the measured tPCB in the laboratory bioassays **must** be coupled with the bioassay results, and that the mean tPCB measurement be associated with the benthic infaunal data.

RESPONSE 3.1-TT-2

Please refer to the response to General Issue 6.C. EPA also performed the analyses using the most synoptic exposure values (see Attachment D.5 of the ERA).

The variability in the field tPCB measurements confounds the construction of dose-response relationships between community structure and tPCBs. Figure D.2-10 presents the results of the 12 individual replicates measurements for tPCBs. Of particular interest, and concern for this assessment endpoint, is that the replicate tPCB concentrations span orders of magnitude. While troubling from a toxicological characterization standpoint – it is a fascinating finding; particularly within the high sand stations (1 – 5). At the December 11th meeting, the Panel questioned the presence of microcapsules of PCBNAPL in the sand as a potential factor in the variability observed in the upstream, high sand stations. EPA has apparently confirmed that PCB micocapsule-NAPLs exist with sand grains, but we were not provided with specific data on locations. A question worth pursuing discussion within the Peer Review is whether these measurements can be taken to represent toxicologically relevant (i.e., bioavailable) concentrations. Porewater measurements at each of these sites might have provided a better means of assessing what is available to benthic infaunal communities, but those data are likely not available.

RESPONSE 3.1-TT-3

EPA agrees that variation in the PCB exposure data introduces uncertainty in concentration-response assessment. However, as indicated in Response 3.1-MO-1 above, significant concentration-response relationships were observed. Please refer also to the responses to General Issues 11.B, 11.D, and 11.G.

EPA agrees that partitioning of PCBs within the coarse-grained sediment (Reach 5A) is atypical, in that organic carbon is not the dominant factor governing the distribution of PCBs (WESTON 2004). Unlike downstream reaches, PCB concentration is not significantly correlated to TOC within Reach 5A (BBL and QEA 2003). Reach 5A sediment, as a whole, appears to deviate from traditional equilibrium partitioning (EqP) theory, which has been shown to adequately describe partitioning in a number of other environments (MacDonald 1994; DiToro and DeRosa 1998). However, explanation for the partitioning behavior in Reach 5A is unlikely to be limited only to the microscale staining of quartz grains with PCBs, which has been observed in a subset of coarse-grained sediment. For example, the validity of the EqP model has been questioned in sediment that contains very low concentrations of organic carbon. In Reach 5A, the organic carbon fraction often falls below the EqP applicability thresholds (Karickhoff 1984; DiToro and DeRosa 1998; DiToro et al. 1991; MacDonald 1994; DeWitt et al. 1992).

EPA has not described the coating of grains as non-aqueous phase liquid (NAPL), and does not believe the coating of grains is the primary reason for the high variability in PCB concentrations observed over small spatial scales. The high spatial variability is also apparent in downstream sediment in which the partitioning of PCBs between adsorbed and aqueous phases appears to be adequately described by EqP principles.

The site-specific bioaccumulation studies conducted by Wright State University using the freshwater oligochaete *Lumbriculus variegatus* suggest that the sediment contaminant concentrations measured at the sediment quality triad locations are toxicologically relevant. Although there is some uncertainty introduced by the short exposure duration (7 days) and the high variability in sediment PCB concentrations, the biota-sediment accumulation factors observed in these tests indicate that the sediment-associated PCBs are bioavailable.

A second factor that will be important to evaluate is whether tPCB concentrations at each of the individual stations represent statistically different concentrations. As noted above, and I believe the Panel all agreed, the data should be used in their entirety, and the central question to ask first is if indeed the stations represent statistically different concentrations. How the assessment of risk throughout this study, and in particular for the infaunal community data is open to re-evaluation based upon the answer to this question. “Eyeballing” the data, it appears to me that that there will be at most only two to three groups of stations that have statistically different PCB concentrations.

RESPONSE 3.1-TT-4

Statistical comparisons cannot be conducted between stations in the toxicity studies because there is only one sediment chemistry sample per station, if only the “most synoptic data” are used (as recommended by the Reviewers). For benthic community data, there are 12 field replicates per station, such that statistical tests (including power analyses) can be performed. These analyses will be included in the revised ERA. EPA does not agree that the assessment of risk depends on specific pairwise comparisons in PCB chemistry among sites. The concentration-response evaluations and determination of an MATC were based on curve-fitting methods that do not require that each concentration be significantly different from others. Please see the response to General Issue 6.D.

At this point, GE's argument in the formal comments to the ERA is compelling: when plots of tPCBs for each replicate are made against taxonomic richness or organism abundance, a dose-response trend is not discernable. The Final ERA must take this into consideration.

RESPONSE 3.1-TT-5

In their presentation, GE acknowledged that there are statistically significant differences in benthic community structure in coarse-grained sediment that are related to PCB concentration; however, GE stated that these differences are not ecologically significant due to effect size. In the January 2004 presentation to the Reviewers, GE presented plots of richness and abundance against tPCB concentration using the replicate data. These relationships were also presented in the ERA (Figures D.3-27 and D.3-28). The main difference in the two presentations is that the EPA analysis considered all the data, including the reference locations, whereas GE excluded the reference stations. The rationale for exclusion of the reference station data was not provided by GE, and EPA believes that the presentation of all data is consistent with the Reviewers' recommendation to use the existing data in their entirety. EPA believes that the data illustrate a significant alteration in benthic assemblages that is related to PCB concentration, and that the response is non-linear. The non-linear nature of the response (i.e., threshold response), combined with the variability in PCB exposure concentrations and natural biological variability, explains the lack of clear concentration-response across the limited range of exposures considered by GE. Please refer also to the responses to General Issues 11.A, 11.B, and 11.G

EPA (2000a; Section 1.3.7.10.2) evaluated the concordance between sediment toxicity and benthic community results for both highly contaminated and moderately contaminated sediment at several northern U.S. sites. Laboratory sediment toxicity tests "better identified chemical contamination in sediment compared to many of the commonly used measures of benthic invertebrate community composition." Accordingly, the less clear response of benthic invertebrate samples is not unusual.

For the moment then, I take position that use of the median value is not appropriate for assessing community data, nor for assessing bioassay results. The ERA presents bioassay results, and subsequently estimates MATCs, using median bulk sediment concentrations (see the presentation in Table D.3-2). I haven't seen a compelling argument that this is appropriate: bioassay results should be based upon the measurements taken at the end of the test – not on a median calculated over all of the bioassay data sets.

RESPONSE 3.1-TT-6

The "most synoptic" data were considered and the analyses presented in Attachment C.5 for toxicity test endpoints. Please refer to the response to General Issue 6.C.

Bedded Sediment and In Situ Bioassays

Each of these tests is discussed below.

The in-lab *Hyalella azteca* tests are very solid data – there is excellent concordance between the results seen at 28, 35, and 42 day survival, the number of young produced, and overall weight gains. I agree with Dr. LaPoint that this is perhaps the most important data set produced within the benthic studies. Dose response is clearly demonstrated for tPCBs vs. 42-day survival and reproduction. One thing that was never made clear in the ERA (or the Work Plan or the EVS report) was why there are three separate time endpoints for this bioassay, and it would be helpful to make a clear explanation. As a corollary to that, in my opinion the only endpoint that should then be used for the MATC calculation is the 42-day survival and reproduction.

RESPONSE 3.1-TT-7

The use of three separate time endpoints is described in the test protocol document cited both in the ERA and in the principal investigator report (EVS 2003). EPA (2000a) describes measurement of the following endpoints during the 42-d *Hyalella azteca* sediment toxicity test: survival (Day 28, 35 and 42), growth (Day 28 and 42), and reproduction (young per female between Day 28 and 42).

The rationale for the separate measurements derives from the test methodology (Ingersoll et al. 1998) that formed the basis for the EPA protocol. In Ingersoll et al. (1998), a series of tests was conducted (42 and 49-d durations) with 7-d to 8-d-old amphipods to investigate techniques for measuring reproductive effects. The 42-d test consisted of an initial 28-d sediment exposure (consistent with the established methods of Ingersoll et al. [1996]) that was intended to run until shortly before release of the first brood of offspring. This initial exposure was followed by an additional 14-d period in which amphipods were held in water to monitor reproduction. The amphipods were expected to be in amplexus (onset of reproduction) between Days 21 and 28, and then release their first brood between Day 28 and Day 42 (Ingersoll et al. 1998).

Measurement of amphipod survival and growth at Day 28 allows for comparison of results from other studies that used a 28-d exposure (i.e., following the Ingersoll et al. [1996] method), and also allows for assessment of these endpoints prior to onset of reproduction. Surviving amphipods from the remaining replicates are transferred from their sediment exposure to their water-only exposure chambers for an additional 14-d period. Both Ingersoll et al. (1998) and EPA (2000a) specify survival and reproduction at Day 35 as test endpoints. The additional measurement was included in both protocols as a weekly monitoring event to assess reproduction and to ensure that adult survival was still high and that test conditions remained acceptable. The Day 35 data also provide useful information about delayed onset of reproduction. Ingersoll et al. (1998) appear to have selected the 42-d exposure duration based on results obtained from a 49-d water-only feeding study with *H. azteca*. All three endpoints (survival, growth and reproduction) are assessed at test termination on Day 42.

EPA agrees that only the 42-day survival measurement should be used in the MATC calculation. Because the Day 42 reproduction endpoints were acceptable for all treatments, they should supersede the Day 35 results. This redundancy will be removed from the MATC calculations in the revised ERA.

However, EPA believes that the chronic tests with endpoints for survival (42-day), growth (28-day), and reproduction (28 to 42 day) are equally important and provide relatively independent measures of contaminant stress and should all be considered in the MATC derivation. Please refer also to the response to General Issue 6.B.

The in-lab *Chironomus tentans* is also a good test, although the fact that only the control and reference sediments showed any survival – and weight gains. This limits the utility of these data in calculating an MATC – although the argument is effectively made here that the NOEL for this test is 0.3 mg/kg tPCBs.

RESPONSE 3.1-TT-8

The Reviewer has not interpreted the test results correctly; some treatments other than the control and reference sediment samples showed survival and/or weight gain. Figure D.3-4 of the ERA illustrates that survival of some chironomids was observed at Stations 4 and 8 at "synoptic" concentrations of 9 and 72 mg/kg tPCB, respectively; however, the survival rate was low (<10%). A similar response was observed for the growth endpoint, as shown in Figure D.3-5 of the ERA. EPA believes that these intermediate responses, combined with the lack of response in reference sediment and 100% mortality observed at the highest concentration tested (213 mg/kg), demonstrate that a concentration-response relationship exists and therefore, that the data are appropriate for use in the MATC calculation.

The conduct and results of the in-situ bioassays appears to be well-thought out and executed. The design and results are in general useful – although I do not believe that using *Daphnia magna* as a test organism for sediment-contact is appropriate. That is a water-borne exposure pathway; the Panel was generally in concurrence on this point.

RESPONSE 3.1-TT-9

Please refer to the response to General Issue 11.E for discussion of the relevance of daphnid test endpoints.

Accepting that for the moment, the greater issue is presenting the results of the tests against "median" tPCB values over all the in-situ exposures. For example, Table 16 of the EVS presents the results for tPCBs for the 48 hour, 7 day, and 10 day exposures. In Table 17, the results of 48-hour sediment exposures are given. Using the results from Table 16, those data do not show a dose-response for *H. azteca* survival. In fact, at the highest concentration measured in sample 031 (522 mg/kg tPCBs), no response was observed even though a response was observed at lower tPCB concentrations. Yet, in Table D.3-2, and in calculation of the MATC, a median bulk tPCB concentration of 77 mg/kg is presented. This looks to me like a "mix and match" of data that are convenient to a finding of risk, and do not follow accepted scientific methods. This becomes especially acute (no pun intended) in calculation of the final MATC. The "mix and match" continues with the presentation of responses to the 10 day in situ exposures. Table 16 shows the 10-day tPCB sediment concentration for Station 428 to be 1.4 mg/kg, with an approximate 50% mortality response for *H. azteca*. No effect is shown at Station 019 at 14 mg/kg. Table D.3-2, however, shows a median tPCB concentration of 7.3 mg/kg tPCBs.

RESPONSE 3.1-TT-10

As discussed above, EPA did not consider only the median exposure data in the MATC calculation; both the median and most synoptic data were considered and presented in the ERA (please refer to the response to General Issue 6.C). Accordingly, EPA does not agree that there was any mixing of data that would violate accepted scientific practice. Indeed, the use of the median and most synoptic data in separate calculations provides a transparent presentation of the MATC derivations using both exposure data processing methods. Inclusion of both types of analyses, however, may have led to the confusion on the part of the Reviewers because the resulting comparison of both measures of exposure for the multiple endpoints measured results in a complex evaluation and discussion.

Although it is true that some individual test endpoints indicated a stronger concentration-response relationship when the median measure of exposure was used, there are also cases where the use of the median resulted in a poorer concentration-response relationship, relative to the most synoptic analysis. Rather than choose one method to the exclusion of other, EPA presented both approaches so that Reviewers could evaluate the two approaches. To this end, Figures D.3-1 through D.3-11 in the ERA clearly labeled the graphs with both sets of PCB exposure calculations.

EPA agrees that the concentration-response for 48-hour *Hyalella azteca* is poor relative to the concentration-response demonstrated for other toxicity endpoints evaluated. As indicated in the response to General Issue 6.D, the decision rules established for determining concentration-response in benthic community endpoints would indicate that this endpoint should be excluded from MATC calculations. EPA notes that a lack of clear concentration-response for *Hyalella* in that particular test does not call into question other toxicity results, because the duration of exposure (48 hours) was not sufficient to allow maximum bioaccumulation or toxicity to occur. Longer-term exposures (both the 10-day in situ test and the 42-day laboratory test) to *Hyalella* resulted in large adverse responses at the highest tPCB exposures tested. As identified in EPA (2000a), short-term exposures that measure survival may not be adequate to evaluate effects of sediment contaminants, which is why longer test durations (and subchronic endpoints) were added to the most recent edition of the testing manual.

This continues over other data points and in my opinion is not appropriate. Results must be paired with the tPCBs that are measured for that test. Otherwise, it appears as if these data are skewed in a way that would point toward risk, instead of letting the data tell us if there is in fact risk present. This absolutely has to be the case for calculation of the MATC.

RESPONSE 3.1-TT-11

EPA presented the concentration-response analyses two different ways, including the method preferred by the Reviewers (see Responses 3.1-TT-2 and 3.1-TT-6 above). Moreover, the choice of data processing methods was not based on a desire to bias the risk conclusion, but rather to acknowledge the uncertainties inherent in the exposure assessment, and provided a bounding analysis using two estimates of exposure. In response to the comments of several Reviewers (General Issue 6.C), however, the calculation of the MATCs using the most synoptic data will be emphasized in the revised ERA.

I want to make it very clear that I believe there is evidence within these data for PCB impacts on benthic infauna – but not within the way it is currently presented.

Benthic Infauna Community Structure

Based on the Panel discussion, I am in agreement that other indicies may be more appropriate, and should be explored. Dr. LaPoint indicated that Shannon-Weiner is not an appropriate index for systems where the ecology is dominated by few species, and recommend Simpson's dominance index. I do not believe that SW should be discarded, but that Simpson's may be more appropriate given the low number of taxa and the confounding course vs. fine-grained substrate. Dr. Oris noted that based on the SW index, there is a distinct response difference between fine grained and course grained sediments. Fine-grained species appear less sensitive than course grained species. I have reviewed that and concur; this needs to be explored further and discussed.

RESPONSE 3.1-TT-12

EPA will provide analyses in the revised ERA for both the Simpson's and Shannon-Wiener diversity indices (please refer to the response to General Issue 11.A). EPA will also discuss response differences related to particle size.

The ERA should do a better job of explaining why in the Work Plan soft depositional sediments were targeted for collection, but in the end six non-depositional (i.e., sand) sites were selected for sampling. Sampling of benthos differed somewhat for upstream coarse grained (shallow water) versus downstream fine-grained (from boat with fauna collected along shore therefore larger spatial separation in latter 10-20 m). This should be accounted for in the discussion of potential uncertainty. Otherwise, I note that the collection of infauna methods, enumeration, and calculation of indicies all appear to be appropriately calculated and represented.

RESPONSE 3.1-TT-13

The term "soft" in the Work Plan was used to distinguish the proposed sampling sites from rock or cobble bottom that could not be sampled effectively with a grab sampler. Very little of the latter substratum occurs in the PSA. Because the benthic study was intended to be representative of the range of benthic communities inhabiting the PSA (except the rock or cobble bottom areas), sampling at sand sites was envisioned at the Work Plan development stage, as can be seen from the locations of the proposed sampling stations, several of which were proposed in areas of the river that are dominated by coarse-grained sediment. EPA agrees that the use of the term "depositional" in the Work Plan may have led to confusion regarding the intent of the study. Additional explanation will be provided in the revised ERA. Please refer to the response to General Issue 11.F.

The overarching issue with the benthic data is the variability of tPCBs relative to the community analysis. As noted above, I suspect that when statistical analyses are applied to the within station tPCB variability, we may not see very much difference in these stations. This is not accounted for in the ERA, or in the subsequent re-analyses presented by EPA in response to question TL12 (EPA Response to Peer Review Questions dated 12/11/03). I remain in agreement with two points in the evaluation of benthic community structure prepared by Scott Cooper of UCSB: (1) variability in tPCBs must be accounted for in this evaluation and (2) habitat factors need to be better accounted for in the benthic analysis.

RESPONSE 3.1-TT-14

EPA believes that it has appropriately considered variability in tPCB concentrations in the evaluation, both in the ERA and in the additional materials supplied to the Reviewers in December 2003. However, based upon the comments from the Reviewers, additional statistical analyses to further explore the variability in tPCB concentrations will be presented in the revised ERA. Habitat factors have already been incorporated in the ERA; however, further analyses conducted by both GE and EPA will be considered in the revised ERA. Please also refer to the responses to General Issues 11.A and 11.B.

Finally, there is no synthesis in the ERA between the bioassay data and the benthic infauna community data. The ERA determines an overall high toxicity for stations 7 and 8 based upon lab and in situ bioassay results, but also shows no community effects at the same stations. This contradiction needs more exploration and explanation than is currently given.

RESPONSE 3.1-TT-15

When the PCB concentrations measured synoptic with the two studies are considered, there is no contradiction in the findings of the two studies. The evaluation of the differences in results of toxicity testing vs. the benthic community metrics at Stations 7 and 8 must consider the differences in PCB exposures observed by appropriately matching effects data with exposure data. The PCB concentrations measured synoptic with the benthic community samples were much lower than those measured in conjunction with the toxicity tests. The toxicity responses at these two stations were associated with synoptic PCB concentrations (laboratory, 48-h in situ, 10-d in situ, respectively) of 213, 139, and 52.3 mg/kg tPCB (Station 7), and 72, 522, and 112 mg/kg tPCB (Station 8). In contrast, the median concentrations in the benthic community grab samples were only 4.3 mg/kg and 14.1 mg/kg, respectively. Therefore, using the most synoptic concentrations, similarly significant responses should not be expected in the benthic community metrics.

Stations 7 and 8 exhibited clear indications of toxicity at the higher concentrations measured synoptic with the in situ toxicity testing, and modest indications of impairment at the lower concentrations measured synoptic with the benthic community sampling. These findings are in concordance. EPA recognizes the potential confusion arising from the different chemistry results for each component of the triad, and will provide a more coherent integration of the lines of evidence in the revised ERA. Please also refer to the response to General Issue 11.C.

Tissue Analyses

As noted above, benthic tissue analyses were originally intended for contributing data to the AQUATOX model, and appear to have been brought into the ERA post priori. Tissue concentrations can be an appropriate measurement endpoint – but it would be useful to have a discussion of what species were in the D-net collections that comprise “predators and shredders”.

If I read the RFI correctly, historically, stoneflies and caddisflies have constituted these groups collected by similar methods. From an exposure standpoint – this is important.

RESPONSE 3.1-TT-16

EPA agrees that tissue analyses provide a useful addition to the sediment quality triad in that they confirm whether COCs are bioavailable to species of concern, and to what degree. It was envisioned in the beginning of the study to use the data in such a way, as well as to support the bioaccumulation model.

Detailed taxonomy was not conducted on the D-net samples because of the large number of organisms needed to achieve the requirements for minimum mass for laboratory chemistry analyses. Sorting of taxa into the predator and shredder groups was conducted by Dr. Bruce Grantham (Lotic Inc.). Based on the sampling method applied, however, it is reasonable to assume that the assemblages were qualitatively similar to the nearby grab samples for which detailed taxonomic evaluations were performed. On this basis and given the collection method, the collected biota would have represented a combination of both epifaunal and infaunal organisms.

The D-net samples were not dominated by stoneflies and caddisflies. These taxa have been collected in some reaches of the Housatonic River, predominatly in faster-flowing reaches of the river. Stoneflies, dobsonflies, and caddisflies have been sampled since 1978 in West Cornwall, CT, but this cobble/riffle habitat is not similar to the warmwater, relatively fine-grained habitat found in the PSA. Based on inference from the benthic grab taxonomy, and screening taxonomic evaluations by Dr. Grantham in the field, the PSA D-net samples likely contained a significant percentage of chironomid larvae and oligochaetes, some caddisflies, and some odonate larvae, but few if any stoneflies. Other taxa, such as gastropods, leeches, and bivalves were more common in downstream PSA stations, and may have contributed significantly to the D-net sample biomass.

MATC Calculation

The MATCs may not appropriately calculated for the following reasons: (1) the response data must be paired with tPCBs measured for that test, (2) only dose/response tests should be included, (3) only one test endpoint should be incorporated into the MATC (i.e., not include 28, 35 and 42 day endpoints), and (4) the MATC should be based upon chronic endpoints – not acute.

RESPONSE 3.1-TT-17

These recommendations are discussed in the responses to General Issues 6.C, 6.D, 6.B, and 6.A, respectively, and will be addressed in the revised ERA.

3.1(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria, based on accepted scientific practices, and incorporated, appropriately in the ERA?

Valery Forbes:

No GE studies performed. GE's reanalysis of benthic community structure is a relevant contribution and should be incorporated.

RESPONSE 3.1-VF-2

Please refer to the response to General Issue 11.A.

Thomas W. La Point:

Yes.

James T. Oris:

No GE studies are presented in the ERA for invertebrates. However, the EPA needs to address concerns raised by GE's reanalysis of data.

RESPONSE 3.1-JO-1

Please refer to the response to General Issue 11.A.

Mary Ann Ottinger:

There were a number of studies conducted by GE over the course of the pre-ERA and ERA. These studies have been included in the ERA, along with discussion of analyses and uncertainties, reanalyses in some cases, and inclusion of the data when appropriate. However, there were no GE studies conducted for this receptor.

Brad Sample:

Although GE did not generate any data for benthic invertebrates, they did perform additional analyses of the community data in relation to PCB concentrations. These multiple regression analyses suggest that PCBs may account for 1% – 9% of total variation in community metrics (data presented in January meeting – analyses do not appear to be represented in GEs comments dated September 29, 2003). From the limited description provided at the January meeting, these analyses appear to be suitable and appropriate. I recommend that GE provide the complete analyses and results to EPA for consideration and inclusion in the ERA.

RESPONSE 3.1-BS-1

Please refer to the response to General Issue 11.A.

Ralph G. Stahl:

There were no GE studies conducted for benthic invertebrates that were listed in the ERA.

Timothy Thompson:

No GE studies were presented to us for consideration for this endpoint. However, the GE reanalysis, and the appendix presented by Dr. Scott Cooper, are appropriate and germane for consideration in the final ERA.

RESPONSE 3.1-TT-18

Please refer to the response to General Issue 11.A.

3.1(c) Were the estimates of exposure appropriate under the evaluation criteria, and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?

Valery Forbes:

Given the extremely high spatial and temporal variability in sediment PCB concentrations (and to some extent other COCs), it is unfortunate that a number of the chemical measurements could not be easily matched with toxicity and/or community structure information. The difference in sediment concentration trends (stations 4 – 8) between the benthic community samples (sediment PCB concentration declines) and the toxicity station samples (sediment PCB concentration increases) is unfortunate and does not increase the clarity of interpretation.

The laboratory toxicity tests should use the most synoptic sediment concentrations for estimating exposure whereas for field community structure it is possible to include paired sediment concentrations from same sites/samples.

RESPONSE 3.1-VF-3

Please refer to the response to General Issue 11.B for discussion of variability of PCB concentrations and implications for risk characterization.

Please refer to the response to General Issue 6.C regarding use of synoptic data.

Thomas W. La Point:

By-and-large, yes. However, there is still uncertainty added to the results by mixing sediments collected at one site into one "batch." In a synoptic approach, it would be better to determine the variance in PCB concentration from distinct samples and exposing laboratory organisms to the various field-assessed range of concentrations.

RESPONSE 3.1-TL-1

The mixing of sediment as described above was conducted only for the sediment toxicity tests to achieve the necessary sediment volumes to conduct the test; the benthic community chemistry values came from small sub-cores collected from individual Ponar grabs, and were not composited. EPA agrees that exposure to a range of sediment PCB concentrations representative of the range of concentrations observed in the field was important. This concept was integral to the original design of the Triad study. Potential sampling locations with varying concentrations throughout the study area were identified from the historical data set. These locations were resampled twice in time with single grab samples to identify the final station locations. At the time of study design and implementation, the small scale variability was not expected or observed. In addition, there were practical limitations to the number of exposure treatments that could be accommodated in the toxicity laboratory; this limited the total number of sampling locations included in the study. Testing of multiple sediment samples (which may have resulted in multiple concentrations) at each location would have required a commensurate reduction in the number of replicates per test, thus reducing the statistical power and increasing the uncertainty for important test endpoints. Furthermore, even had it been known at the time of study implementation, the small-scale variability in PCB concentrations would have been a difficult factor to control for in an experimental design, because repeated measurement of sediment subsamples, even from the same grab, may have yielded differing PCB chemistry results.

James T. Oris:

Estimates of exposure and the COC refinement was appropriate. However, there seems to be some discrepancy between TOC and PCB concentration relationships with no explanations.

RESPONSE 3.1-JO-2

There is substantial information in the RFI (BBL and QEA 2003) on the subject of the relationship between TOC and PCB concentrations in the PSA. This information was considered in the ERA, and will be referenced to a greater degree in the revised ERA. Please refer also to the response to General Issue 11.

In addition I am confused about the relationships between chironomid abundance versus taxonomic diversity related to substrate (fine, coarse) and tPBC concentration. These relationships appear to be opposite of one another (slides 36 and 38 in invertebrate presentation from document presentation meeting). This difference indicates that general conclusions about invertebrate toxicity in fine versus coarse sediment related to tPCB cannot be made. Sensitive species (chironomids) seem to have higher rates of toxicity in fine sediments, but species diversity seems to be more sensitive to tPCBs in coarse sediments. This will require further discussion and deliberation.

RESPONSE 3.1-JO-3

Please refer to the responses to General Issues 11.A, 11.B, 11.D, and 11.G and Response 3.1-RS-8. EPA does not agree with the conclusion that chironomids have higher rates of toxicity in the finer sediment. In EPA 2000a (Section 14.4.3), the relative tolerance and lack of impact on measures of effect of *Hyalella* to a large gradient in grain size is noted, and, if anything, a weak correlation of decreased survival with increased grain size (Ingersoll et al. 1998) was documented. Although it is true that the fine-grained stations 7 and 8 consistently exhibited strong toxicological responses, EPA believes that a principal reason for that finding is that the PCB concentrations measured synoptic with the toxicity tests were, on average, much higher than at the coarser-grained locations.

Because coarser-grained stations were not tested for toxicity at the same PCB concentrations as the finer-grained stations, the issue of the comparative sensitivity of chironomids in different substrates cannot be directly assessed considering the toxicity data alone. As described in the response to General Issue 11.C, EPA will provide a more thorough analysis of the concordance among triad components in the revised ERA.

Mary Ann Ottinger:

The assessments appear to be appropriate. However, organochlorine pesticides were eliminated from the ERA due to apparent contradictions in the laboratory analyses leading to reducing the estimates (reference to laboratory interference may be found in Volume 2. 6.2). This means that impacts due to effects of mixtures were not considered, even with potential spotty exposure as discussed below.

RESPONSE 3.1-MO-2

Although it is true that some of the pesticides detected in tissue samples may be artifacts due to interference in the laboratory, this was not the primary reason for exclusion of pesticides from the benthic invertebrate ERA. Section D.2.1.1 of the ERA describes the receptor-specific COPC screening, and includes an evaluation specifically for pesticides. This additional screen "considered the frequency of detections in available tissue chemistry data for these contaminants (which were not evaluated in the Pre-ERA), combined with a preliminary conservative ecotoxicity screen using conservative tissue concentrations."

An examination of the analytical results for invertebrate tissue indicated that concentrations for most of the pesticides of concern (except dieldrin, endosulfan, and hexachlorobenzene) were below detection limits. For detected pesticides, measured concentrations were below tissue concentration effects thresholds compiled by Jarvinen and Ankley (1998). Based on these considerations, the entire suite of organochlorine pesticides was eliminated from further consideration in the invertebrate portion of the ERA. Because pesticide risks to invertebrates were considered negligible, an assessment of interactions between pesticides and PCBs was not warranted.

The estimates of exposure were based on a measurements of samples collected in surveys of aquatic, sediment, and soil samples, assessing the results of these analyses, and then refining the longer list to COCs based on levels and persistence of chemicals (i.e., likely to result in impact to indigenous organisms). This is a reasonable approach that meets EPA criteria. The only uncertainty remaining is that in most cases, pesticides and other contaminants are generally not considered. Therefore, areas of the PSA that include or border agricultural fields or golf course are likely to contribute pesticide load to the aquatic environment; especially during runoff events, which were not considered in the models. The patchy nature of this type of exposure also leads to uncertainty and is in the realm of factors that are present, but hopefully do not result in substantial impact upon the wildlife in the ecosystem.

RESPONSE 3.1-MO-3

Please refer to Response 3.1-MO-2 above. The available tissue data do not indicate pesticide loadings that would result in risk. In addition, sediment sampling in the PSA was extensive, and 10% of the samples were analyzed for Appendix IX contaminants, including pesticides. Although it is theoretically possible that localized and unsampled areas exhibit higher pesticide concentrations (or other COPCs) than the analytical results obtained from the existing sampling program, there is no evidence that this is actually the case. The COCs that remained following the receptor-specific screen were explicitly evaluated in the ERA.

In the case of many of the models, the relationship between TEF, TEQ, and LD50 data from the literature are not implicitly obvious. This is due, in part to the complexity of the data and consequently the necessary sophistication of the models used to estimate risk. However, as pointed out above, there needs to be a clear summary linking the chemicals to the species, their sensitivity, and the resulting potential risk in areas of the PSA and downstream. This would facilitate decision-making regarding future action. For example, understanding the association of data shown in Figures 7.4-1, 7.4-2 and linking these data to interpretations from the field studies are somewhat difficult and confusing.

RESPONSE 3.1-MO-4

The TEF and TEQ models do not apply to the benthic invertebrate ERA. The studies used to derive TEF values are based upon avian, mammalian, and fish toxicity. Benthic invertebrates are not known to exhibit toxicity via the Ah-receptor mechanism that is responsible for effects in vertebrates; therefore, toxic equivalency models were not considered appropriate exposure methods for invertebrates.

Brad Sample:

The sediment concentration data were highly variable over very small spatial scales. This posed a significant difficulty in associating field and laboratory data to appropriate exposure data. EPA generally did a good job of describing the variability and attempting to account for it in their analysis. There were however some problems with the strength of the association of exposure data to effects data that add uncertainty and somewhat reduce the overall strength of the analyses as presented.

RESPONSE 3.1-BS-2

Please refer to the responses to General Issues 11.A, 11.B, and 11.D, regarding variability (in exposure and effects) and the linkage to concentration-response evaluations.

It appears that the exposure data used to determine the dose-response relationship for tPCB was the median concentration observed among all sediment samples collected in 1999. I believe that this is an error. The bioassay results relate best to the sediment concentrations to which the test animals were exposed – this should be the ‘most synoptic’ sample concentrations. The dose-response-relationships used for the laboratory bioassays, and all subsequent analyses that use these results, should be revised to take this change into account.

RESPONSE 3.1-BS-3

Please refer to the response to General Issue 6.C. Two approaches were evaluated in the ERA, both the relationship between effects and the median concentration of samples associated with the study site, and the effects and the "most synoptic" sediment concentration, as recommended by the Reviewer.

From the discussion at the January meeting, it appears that sampling of benthic invertebrates differed between the shallower upstream areas and the deeper downstream areas. This difference is not immediately apparent in the ERA report. Whereas in the upstream areas, invertebrates were collected in near proximity to sediment samples, this was not possible in the downstream areas. Sediments in the downstream areas were collected by boat. Benthic invertebrate samples were collected from locations near the shore – some 10-20 m distant from the sediment samples. This spatial discontinuity needs to be made clear in the report and described in the uncertainties section. The key question that needs to be looked into is can a gradient exposure be inferred based on these data? Given the high spatial variability, these downstream data may not be adequate.

RESPONSE 3.1-BS-4

The benthic community sampling was conducted using the same methods at all stations. The only difference was that the grab sampler was deployed by wading in the shallower areas and from an anchored boat in the deeper areas where wading was not practical. No location had water depths exceeding approximately 4 feet; therefore, it was possible to ensure that the area of bottom encompassing the 12 replicate grabs was approximately the same, regardless of which sampling method was used. Sediment chemistry samples were removed from each of the grabs and thus are as synoptic as practically possible. Samples for tissue analysis were collected in the general vicinity of each grab sampling location. Although there was generally less separation between the community grab samples and the tissue analysis D-net samples at the upstream stations, the difference in distance between the samples in the upstream and downstream locations was not large (approximately 3 to 6 meters vs. 10 to 20 meters). Please refer also to the response to General Issue 11.F.

Ralph G. Stahl:

As noted above, it appears that the most synoptic sediment exposure values for tPCBs were not used in the ERA. Therefore, the exposure estimates for benthic invertebrates does not meet the evaluation criteria of consistent and objective.

RESPONSE 3.1-RS-4

Please refer to the response to General Issue 6.C. The most synoptic sediment exposure values were presented and analyzed with the effects data in Attachment D.5 of the ERA.

Timothy Thompson:

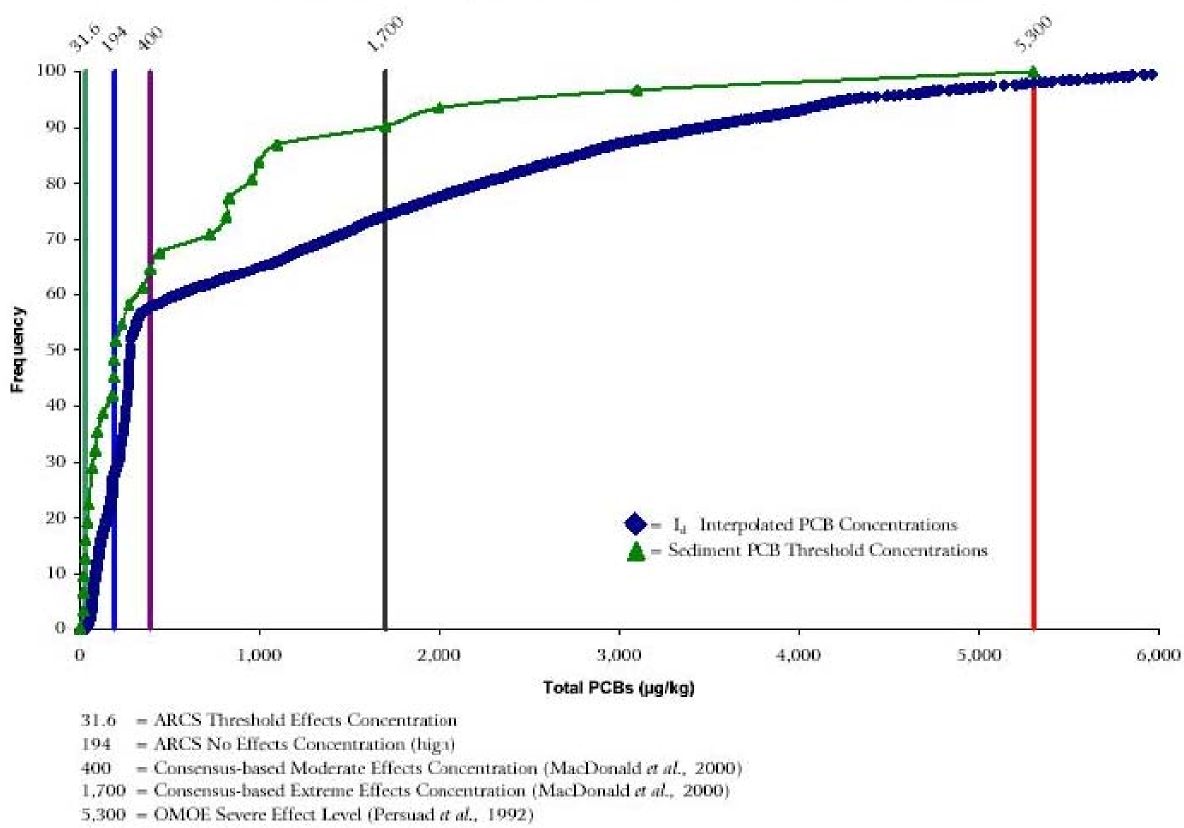
See comments under (a), above.

The ERA in made good use of probability distributions – and in particular the distribution of COC concentrations and the probability of exceeding the HQ (or MATC). This was not done in the case of benthic infauna, and should be conducted for each of the COC within the PSA. I presented an example that was used on the Lower Fox River and Green Bay PCB ERA, which is included below. These are simple frequency distributions of both measured, and GIS-base map interpolated tPCB concentrations. Something equivalent constructed for the PSA (and for the Connecticut reaches below the PSA) would provide a good visual presentation that is consistent with that done for other receptors in this ERA.

RESPONSE 3.1-TT-19

Please refer to the response to General Issue 1.C. The revised ERA will provide cumulative frequency plots of sediment exposures, against which multiple effects thresholds are compared. The effects thresholds will include values based on both acute and chronic responses (in accordance with the response to General Issue 6.A), and will consider multiple effect sizes (see Response 3.1-JO-4 below.)

Figure 6-88 Green Bay Zone 2 Sediment PCB Frequency Distribution



3.1(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria?

Valery Forbes:

I question the use of Daphnia and Ceriodaphnia as appropriate benthic invertebrate test species. It would have been better to use another infaunal or epifaunal temperate invertebrate.

RESPONSE 3.1-VF-4

Please refer to the response to General Issue 11.E.

With regard to differences in the relationship between taxonomic diversity and sediment PCB in fine - versus coarse-grained habitats, it could be that the substrate difference is explained by differences in taxonomic composition between fine and coarse sites or that there are differences in PCB bioavailability (e.g., less bioavailable in fine -grained sediments) that could explain these differences.

RESPONSE 3.1-VF-5

Please refer to the responses to General Issues 11.A, 11.D, and 11.G. In the revised ERA, EPA will provide an expanded discussion of the relationship between concentration-response relationships and substrate/habitat conditions.

Sampling of benthos in the field differed somewhat for upstream coarse grained (wading in shallow water) versus downstream fine-grained (from boat with fauna collected along shore therefore larger spatial separation in latter 10-20 m). Whereas this may have been unavoidable, the differences should be mentioned in the discussion of fine- vs. coarse grained site differences.

RESPONSE 3.1-VF-6

Please refer to the response to General Issue 11.F.

It seems that the MATCs are ultimately based on only two species with multiple (nonindependent) response endpoints, and this should be rectified. With regard to deriving MATCs, it is recommended that acute and chronic test endpoints be separated, that only one endpoint be used per species (could be lowest or could be geometric mean), that only the most synoptic data are used as measures of exposure, that only those tests that displayed a clear concentration-response relationship be used, that only sediment-relevant test species be used, that all of the available test species be used (i.e., not just the lowest 6 values), and that if the derived MATC is equal to or lower than the concentration at reference sites the value should be truncated at the reference concentration.

RESPONSE 3.1-VF-7

Please refer to the response to General Issue 6 regarding application of rules for MATC derivation. EPA agrees that the MATC for benthic invertebrates was influenced significantly by the responses observed in *Hyalella* and *Chironomus*. However, this is not surprising, given that these species were the only two for which chronic sublethal test endpoints were evaluated. The remaining test endpoints included a test organism (i.e., *Lumbriculus*) that is known to be relatively insensitive in toxicity tests and is used for that reason to measure bioaccumulation, and some test organisms with very brief exposures evaluating the potential for acute responses (i.e., 48-h survival for *Hyalella*, *Chironomus*, *Daphnia*, and *Lumbriculus*). Nevertheless, the responses observed in the in situ tests (particularly the 10-day exposures) were sufficient to confirm that adverse effects occurred both in the field and the lab, and also occurred to organisms other than *Hyalella* and *Chironomus* (e.g., *Daphnia, Ceriodaphnia*).

Thomas W. La Point:

Yes, particularly the toxicity endpoints using surrogate species.

James T. Oris:

Development of MATC values seems somewhat subjective. However, it was good that multiple endpoints (EC20 and EC50) were evaluated. Would it not be better to use cumulative frequency distributions of EC20's or EC50's and choose MATC based on a measure of central tendency from the distribution rather than the subjective description used in the document?

RESPONSE 3.1-JO-4

Please refer to the response to General Issue 6 regarding rules for MATC derivation.

Mary Ann Ottinger:

The effects metrics were a compilation based on the surveys and data collected as part of the pre-ERA and ERA. Relatively traditional wildlife assessment endpoints were used, including survival and reproductive measures. These data are valuable, but not always sensitive indices of exposure. In the case of measures of the benthic invertebrate community, the metrics appeared appropriate, overall. One study observed that survival was impaired in one species (*Chironomus*) tested in 4 locations. Due to the varied conditions, sediment composition, and changing aquatic conditions, effects metrics for the benthic invertebrate community are complex. Given the complexity of this class of receptor, more conservative MATCs are appropriate.

RESPONSE 3.1-MO-5

In the revised ERA, EPA will recalculate the MATC following the guidelines discussed in the response to General Issue 6.

Brad Sample:

There are some issues with the effects metrics as presented that need to be addressed:

- As described in 1.c - site-specific effect levels were based on the median sediment concentrations. I believe that this was inappropriate. The ‘most synoptic’ values should be used instead. This will result in a change in the effects levels and will need to be carried through the rest of the section.

RESPONSE 3.1-BS-5

Please refer to the response to General Issue 6.C.

- The interpretation of the bioassays as presented includes the use of discontinuous dose-response relationships (less severe effects observed among some treatments with higher doses). For example, this is the case for *Hyalella* survival in Figure 3.3-1 based on median PCB concentrations. These relationships call into question the suitability and validity of the dose-response relationship. It may be an indication that other factors or COCs may be contributing. (Note that this if the exposure is based on the most synoptic data, this discontinuous dose-response is not an issue…at least in Fig 3.3-1). If discontinuous dose response relationships are observed, the reasons why this occurred should be evaluated and discussed. Additionally, if the anomalous dose-response values cannot be explain such that they can be excluded from the set of values, these bioassay results should not be used to calculate toxicity values.

RESPONSE 3.1-BS-6

Please refer to the response to General Issues 6.D and 11.C.

- I am not sure that daphnids are suitable for evaluating sediment toxicity. Although they are suitable to look at bioavailability of PCBs, they are only exposed via water and not through ingesting sediment. These data, while useful do not contribute much to the sediment assessment – they may be more suitable for fish assessment.

RESPONSE 3.1-BS-7

Please refer to the response to General Issues 11.E. EPA does not believe that the daphnid toxicity data are applicable to the fish assessment because the mechanism of toxic action for dioxin-like chemicals differs between fish and invertebrates.

Ralph G. Stahl:

The effects metrics were appropriate under the evaluation criteria. As noted above however, the comparisons made between the treated and laboratory controls in the site-specific toxicity studies generally is not standard practice and should be articulated in the ERA.

RESPONSE 3.1-RS-5

EPA believes that the effects assessment follows accepted scientific practice. EPA calculated endpoints based on comparisons to the negative control and the two reference sediment samples, to provide the most complete evaluation of the data set, but then only used the reference sediment comparisons for development of the MATC. In the risk characterization, indications of potential for harm "were standardized to appropriate background conditions (e.g., toxicity endpoints were compared to reference Stations A1 and A3 rather than negative control sediment)". Please refer to Response 3.1-RS-3 above.

Timothy Thompson:

See comments under (a), above.

3.1(e) Were the statistical techniques used clearly described, appropriate, and properly applied for the objectives of the analysis?

Valery Forbes:

The statistical methods seem generally appropriate. However, the ERA could benefit from a better description of the statistical methods used. Enough detail should be presented so that the analyses could be repeated.

RESPONSE 3.1-VF-8

Please refer to the response to General Issue 1.C.

Shannon-Wiener may not be best measure of diversity for the sediments in which a few species dominate (Tom La Point suggested Simpson's index).

RESPONSE 3.1-VF-9

Please refer to the response to General Issue 11.A.

I believe that the concerns raised by GE in response to the reanalysis of the benthic data are important. If a small fraction of the total variability in benthic species abundance can be explained by PCB concentration, despite statistical significance of the regression, this suggests that the role of PCBs in determining benthic community structure may be less important than concluded by EPA.

RESPONSE 3.1-VF-10

Please refer to the response to General Issue 11.A.

I recommend that both effect size and significance are important and should be presented for all experimental results where appropriate. This is true throughout the ERA.

RESPONSE 3.1-VF-11

Please refer to the responses to General Issues 1.C and 11.A.

Thomas W. La Point:

The multidimensional scaling (MDS) approach is similar to a principal components analysis (PCA), as acknowledged in the text (pg. D-67). I would suggest that a log-linear categorical approach (Sokal & Rohlf 1989) also be employed, as the categorical approach to percentage responses (not only toxicity endpoints, but dominant species groups abundances) and PCB concentrations may yield further supportive information concerning community-level responses to PCB exposure. Finally, the data collected are most appropriate for canonical correlation techniques, such as detailed in Tabachnick and Fidell (1996).

RESPONSE 3.1-TL-2

EPA agrees that multivariate approaches may offer additional insights regarding the role of PCBs in affecting benthic invertebrate endpoints. EPA will consider canonical correlation techniques in the revised ERA, including a discussion of benefits and limitations of the statistical method. The findings will be combined with other statistical analyses, including those presented in the ERA, or provided to Reviewers by GE and EPA in December 2003. Due to the complexities introduced by non-linear dose-response relationships, multiple chemical measurements per Triad station, and other factors, it is best to consider multiple lines-of-evidence, rather than seek to identify a single "most appropriate" method.

The use of Shannon-Wiener diversity, H, to describe the benthic invertebrate community must be questioned. There are numerous problems with it: if the species richness is "low" (fewer than 100 species), then H' is a insensitive measure of the relative frequency of species (May 1975). Further criticisms of the use of H' include (Green 1979; Moriarty 1999; pg 242): H' is unaffected if one species replaces another (there is no consideration of the differences in species taxonomy among sites); communities are not "supra-organisms" and should not have a biological meaning to it; any index needs taxonomic identification to species level – not sub-family. The data are present to completely and successfully analyze total abundances, relative frequencies (perhaps employing a dominance index), or discuss a variety of indices, including the index of biotic integrity, IBI.

RESPONSE 3.1-TL-3

Please refer to the response to General Issue 11.A. Based on the Reviewers' comments, the Simpson's diversity index will be added to the revised ERA.

James T. Oris:

The statistical approaches used in the invertebrate section are not always clear and are not fully explained in the document. It is not possible to evaluate whether statistical techniques were properly applied in all cases since insufficient detail is present to make such an evaluation. Analysis of statistical power, when appropriate, is not provided and queries from panelists concerning power analyses were not adequately addressed.

RESPONSE 3.1-JO-5

Please refer to the response to General Issue 1.C. Some statistical power calculations (MSDs from toxicity endpoints) were provided along with the December 2003 responses to initial Reviewer questions, and will be incorporated into the revised ERA.

The approach described in Appendix C.2 for handling non-detects would appear to be appropriate, but it is not clear that the procedure was used in all cases. Sections in both fish and invertebrate assessments refer to Appendix C.2, but then unless I missed something, non-detects in the invertebrate assessment were set at 0.5 of the detection limit and were set to the value of the detection limit in the fish assessment. In both assessments then, it appears that fixed-replacement was chosen (with little evidence to support the decision), but two different replacements were used. There seems to be a disconnect between the two assessments and it is not clear how these values were chosen.

RESPONSE 3.1-JO-6

Neither the fish ERA nor the benthic invertebrate ERA used a default arbitrary fixed replacement method for estimating non-detected concentrations. Consistent with the methods described in Appendix C.2, the first step was to conduct a bounding analysis setting ND=DL and ND=0 to evaluate the sensitivity to detection limit replacement. In some cases, the analyses were insensitive to the choice of replacement method; therefore, setting ND equal to ½ DL was appropriate. EPA agrees that for some of these cases, additional description of the bounding analysis would be helpful and will provide the information in the revised ERA.

In other cases, such as the calculation of TEQ values in fish tissues, the analysis was more sensitive to non-detected values. In the latter cases, the ERA presented risk estimates made using both the upper and lower bounds to explicitly describe the uncertainty associated with detection limits.

Mary Ann Ottinger:

The statistical techniques appear adequate, however there are a multitude of questions regarding the appropriate statistics, given the data collection methods and the necessity to integrate in the grain size and other variables that are inherent in the river system. As such, the statistical analyses should consider spatial and species characteristics in a multivariate approach. The resulting models should be discussed relative to these variables, both in the PSA and downstream.

RESPONSE 3.1-MO-6

Please refer to the response to General Issue 11.

Brad Sample:

In general it appears that the statistical analyses were correct. However, adequate detail in the description of the methods and exact data used are lacking. Consequently a definitive conclusion cannot be made. This is a general global statement for all statistical analyses throughout the report. There is insufficient transparency in relation to the exact methods used and whether the data met the assumptions for the method. I would recommend including a brief description in each section about the statistical methods employed, and possibly include a discussion in Appendix C in which detailed descriptions of the statistical methods, their application, assumptions, and limitations, are provided.

RESPONSE 3.1-BS-8

Please refer to the response to General Issue 1.C.

Ralph G. Stahl:

The statistical techniques were not clearly described and their application to the toxicological assessments should be reviewed by the authors.

RESPONSE 3.1-RS-6

Please refer to the response to General Issue 1.C.

The calculation of the MATC for tPCBs in sediment does not meet the evaluation criteria because it is not based on the most synoptic data collected for the toxicity assessment.

RESPONSE 3.1-RS-7

Please refer to the response to General Issue 6.C.

Timothy Thompson:

See comments under (a), above.

The statistical methods seem generally appropriate. I concur with my colleagues who noted that the concerns raised by GE in response to the reanalysis of the benthic data are important. As noted by Dr. Forbes, "If a small fraction of the total variability in benthic species abundance can be explained by PCB concentration, despite statistical significance of the regression, this suggests that the role of PCBs in determining benthic community structure may be less important than concluded by EPA.". This relates to the re-analysis done by Dr. Scott Cooper, and is relevant.

RESPONSE 3.1-TT-20

Please refer to the response to General Issue 11.A.

ERA could benefit from a better description of the statistical methods used. Enough detail should be presented so that the analyses could be repeated. A general recommendation: is that both effect size and significance are important and should be presented for all experimental results where appropriate. This is true throughout the ERA.

RESPONSE 3.1-TT-21

Please refer to the response to General Issue 1.C.

Shannon-Wiener may not be best measure of diversity for the sediments in which a few species dominate (Dr. Lapoint suggested Simpson's index).

RESPONSE 3.1-TT-22

Please refer to the response to General Issue 11.A.

3.1(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria?

Valery Forbes:

Regarding the multiple regression analysis provided in response to Panel's questions it would seem that the role of PCBs as a major factor influencing the abundance of benthic invertebrates is questionable. Both proportion of variance explained as well as statistical significance need to be taken into account in interpreting these analyses.

RESPONSE 3.1-VF-12

Please refer to the responses to General Issues 1.C and 11. EPA agrees that effect sizes and contribution to total variability should be considered in the ERA. However, EPA also believes that the sources of natural variation, measurement and analytical variability, and the appropriateness of the statistical models are also important considerations for interpreting the results.

The risk terminology used to describe HQs (i.e., definitions of low, moderate and high risk) needs checking for consistency with other COCs as well as with other assessment endpoints throughout the ERA. HQs should be used as rough estimates of relative risk within assessment endpoints. Broad brush order of magnitude differences could be useful indicators of relative risk. Other COCs have HQs greater than one but the contribution of these was downplayed. See figure 4.2. There is a need for greater consistency in the interpretation of HQs exceeding one. Also the magnitude and frequency of exceeding the relevant threshold should be considered. It is essential to point out that for PCBs variability in the HQs reflects variability in the exposure estimates, with a single value representing the effects. For other COCs HQ variability reflects variability in the effects thresholds with a single point estimate for exposure.

RESPONSE 3.1-VF-13

Interpretation of HQs – Please refer to the response to General Issue 7.

Representation of HQ Variability – Please refer to the response to General Issue 1.C. EPA generally agrees with the Reviewer comment and will provide more information and clarity in the revised ERA. Although the HQ graphs do convey some aspects of the HQ distribution (mean, median, range, percentiles), EPA acknowledges that improved representation of risk could be provided through the use of cumulative probability plots and other presentation methods. Furthermore, EPA agrees that the risk terminology for aquatic receptors can be made more parallel with the wildlife sections, considering both the probability and magnitude of different levels of harm (i.e., exceedance probabilities). Additional clarity will be provided for HQ determinations, specifically indication of whether the spread in the HQs derives from variability in exposure (the numerator), variability in effects (the denominator), or both.

Thomas W. La Point:

The degree of risk to aquatic macroinvertebrates is judged to be high, and I strongly concur.

James T. Oris:

The general characterization of risk is supported by the available information and appears to be appropriate. However, there are several areas of needed improvement in the document and these are included below as separate comments.

Risk terminology used within the invertebrate section needs to be revised and brought into concordance with other assessments (e.g., wildlife sections). Based on current terminology, any HQ value less than 1 presents no potential risk. HQ values greater than one present varying levels of risk on what seems to be a scale of HQ=1-3 (low risk), HQ= 3-10 (moderate risk), and HQ>10 (high risk). However, this scale is neither formally defined nor justified in the document, and the scale seems to be arbitrary. Based on the methods of developing benchmarks (non-conservative) and the use of high-levels of effect, all HQ values greater than 1 should indicate significant potential for risk. Levels of HQs that approach 1 may also indicate risk potential. For example, there is little possibility that there is a difference in risk potential between an HQ=0.9 and HQ=1.1, yet under the current scenario, the HQ=0.9 is considered "no risk".

RESPONSE 3.1-JO-7

Please refer to the response to General Issue 7.

HQ values were interpreted in the ERA as providing only an approximate indication of magnitude of risk. The final characterization of risk was made following the weight-of-evidence (WOE) evaluation for each receptor. The HQ graphs in Section 12 were risk communication tools used to support, but not replace, the WOE risk determinations. The patterns and ranges of HQ values are cited in the text, but correspondence between those HQ ranges and risk terminology is coincidental.

Application of HQ values in the risk characterization is wholly deterministic, even though a probabilistic treatment could be used (such as in wildlife assessments). It is strongly suggested that the risk terminology and the risk characterization be conducted using the approach used in the wildlife sections. This approach, defined using exceedance probabilities will be more appropriate.

RESPONSE 3.1-JO-8

Please refer to the response to General Issue 1.C and to Response 3.1-VF-13 above. The use of species sensitivity distributions (SSDs), probabilistic representation of exposure concentrations, and consideration of multiple effect sizes will result in an aquatic endpoint assessment that is more similar to the wildlife assessments.

A comparison of section D.4.2.1.3 (other COCs) and the data (Figs. D.4-1 and D.4.-2) indicates a high potential for risk due to PCBs, but downplays the potential importance of risk due to other COCs. The figures and the description of the HQs in the figures states that many of these COCs had values greater than 1 and less than 10, yet these potential risks are described with judgmental language that downplays the potential risk. Barium, chromium, copper, lead, mercury, silver, and tPAH all had median HQs > 1 at more than one site. Mercury had an upper range HQ of close to 100 at one site. Only median HQs were considered. The potential risks due to other

COCs were downplayed based on the relative values of HQs between PCBs and other COCs. However, I submit that tPAHs and several of the metals listed above (esp. Hg) can serve as additional significant sources of risk. The HQ approach used here cannot separate out relative effects of PCBs and other COCs. Once an HQ is greater than 1, there is a significant potential for risk. To say that other COCs were not important because their HQs were "barely over 1" and "most often less than 10", is not appropriate. In addition, this approach cannot assess additive or synergistic effects that may occur among more than one contaminant (e.g., PAHs + PCB, or Hg + PCB).

RESPONSE 3.1-JO-9

EPA believes that HQs for other COCs were appropriately considered in the ERA, for several reasons:

- The HQs for other COPCs were based on conservative screening SQVs, which have been demonstrated to yield a high percentage of false positives (Chapman and Mann, 1999). This conclusion is supported by the ERA findings, which indicated a site-specific tPCB MATC for benthic invertebrates that is more than an order of magnitude higher than the median literature-based SQV. Therefore, the caveats attached to the use of SQV based HQ screening values, for both PCBs and other COCs, are appropriate and consistent with the sediment quality triad approach.

- The HQs derived from the COPC and COC screening analyses were not used to make definitive conclusions regarding risk, but instead were used to indicate order of magnitude differences across contaminant types. In this context, the HQ values for other COCs were compared to PCBs, and the differences were observed to be large (i.e., two orders of magnitude). EPA believes that such large differences are sufficient to rank PCBs higher in terms of potential toxicity, and agree that the HQs cannot be used to make definitive statement of absolute risk.

- Other COCs with HQ>1 were not eliminated from further consideration based on the HQ results. Other COCs were retained in the ERA, and other lines of evidence, such as TIE results and correlations of chemistry to effects measures, were used to make risk conclusions.

Also please see the response to General Issue 7 that describes the different uses of the term “HQ” throughout the ERA.

Mary Ann Ottinger:

The information from the models was used in a WOE approach to estimate and characterize risk. This approach was reasonable, especially considering the complexity of the ERA. There are numerous uncertainties, which have been discussed above. The approach of using the threshold for COCs derived from available literature using most sensitive and most tolerant species to develop a threshold range [see 2.6-25] should yield a conservative estimate of risk. The conclusion of high impact to populations and declining risk with distance from PSA appears appropriate. The inherent uncertainties must be considered for future decision making relative to remediation activities.

RESPONSE 3.1-MO-7

EPA concurs with the Reviewer's comments.

Brad Sample:

The risk characterization appears to be generally appropriate with two exceptions – one being the derivation of the MATC and two being the application and interpretation of this value.

MATC derivation:

- MATCs were calculated by combining both acute and chronic bioassay results. To allow greater resolution in the nature and magnitude of potential effects, these data should not be combined. Rather, separate MATCs should be calculated based on results from acute and chronic bioassay results.

RESPONSE 3.1-BS-9

EPA will calculate separate effects thresholds for acute and chronic endpoints. Please refer to the response to General Issue 6.A.

- MATCs were calculated based on multiple measurements of the same type of effect for the same test species. The effect of this is to potentially over-represent effect types and receptors in the derivation of the MATC. I recommend using only a single value for each effect type for each test species per MATC (i.e., use only one *H. azteca* survival value). This issue will be partially addressed by differentiating/calculating acute and chronic MATCs.

RESPONSE 3.1-BS-10

EPA agrees and will calculate the MATCs using one of the values for effect type per species. Please refer to the response to General Issue 6.B.

- MATCs were calculated based on sediment-associated test species and water-column test species. I think that this is inappropriate – given questions concerning exposure for the water-column test species, only sediment-associated species should be used.

RESPONSE 3.1-BS-11

EPA believes that there is a valid rationale for including the species in question in the analyses. Please refer to the response to General Issue 11.E.

- To ensure a defendable dose response relationship, test results should be based only on the most synoptic data.

RESPONSE 3.1-BS-12

This analysis was included in the ERA in Attachment D. Please refer to the response to General Issue 6.C.

- Any toxicity value for which a discontinuous dose-response relationship was observed should be excluded.

RESPONSE 3.1-BS-13

EPA will evaluate the dose-response relationship for the test endpoints. Please refer to the response to General Issue 6.D.

- It appears that the MATCs were calculated based on the lowest six toxicity test values. Justification for this is not provided. I would recommend that the MATC be calculated based on all acceptable toxicity values.

RESPONSE 3.1-BS-14

EPA agrees and will include all applicable values subject to the other decision rules. Please refer to the response to General Issue 6.F.

- It is possible that a derived MATC may produce a value that is lower than what is observed at reference locations. If this occurs, the MATC should be truncated to the reference concentration.

RESPONSE 3.1-BS-15

EPA agrees. Please refer to the response to General Issue 6.G.

- Based on the comments above, a set of rules for deriving MATCs can be outlined: 1) separate acute and chronic 2) do not use multiple endpoints for the same receptor, 3) use only most synoptic data, 4) only use those tests that displayed a dose response, 5) use only sediment-relevant receptors 6) do not restrict data to the lowest 6 values – use all. 7) if MATC is equal or lower than concentration at reference – truncate at the reference concentration.

RESPONSE 3.1-BS-16

EPA agrees with the vast majority of these points. Please refer to the response to General Issue 6.

MATC (and other sediment quality value) application:

- Once calculated, MATCs were applied to all sediment samples within the PSA. This is good. However, the results of this analyses were presented as mean (range) HQs for each area evaluated. This misrepresents potential risks. HQs are basically a binary measure – exposure exceeds or it doesn't. Magnitude of exceedance is not directly interpretable – interpretation is a function of what the toxicity measure represents. I think a better way to present these results would be to tabulate the frequency of exceedance. The greater the number of samples that exceed the MATC (or other toxicity value) the greater the risk. This approach can be refined by using two toxicity values (say an acute MATC and a chronic MATC). Screening results can then be viewed categorically – frequency of samples that do not exceed anything (no risk), frequency that exceed only the chronic value (moderate risk), frequency that exceed both acute and chronic values (high risk).

This point by point analyses also would allow for the identification of hotspots and could provide spatial direction for future remedial actions.

RESPONSE 3.1-BS-17

EPA agrees that such an analysis would contribute to the interpretation of risk and will include such an analysis in the revised ERA.

Ralph G. Stahl:

The characterization of risk is not fully supported by the available information. However, the process of characterizing the risk was appropriate under the evaluation criteria.

The uncertainty with the characterization of risk to benthic invertebrates is high. Benthic abundance throughout the PSA was variable and based on one sampling event. Site-specific toxicity tests indicate impacts at two stations in the PSA (7&8), yet there are not similarly significant depressions in benthic abundance at these same locations. The risk of harm to benthic invertebrates is not consistent across the PSA, and appears to be localized to these two stations. There is no clear dose-response from the site-specific toxicity tests which also tends to cloud the evaluation. The results of the TIE, while interesting, are not conclusive with respect to tPCBs causing the observed toxicity.

RESPONSE 3.1-RS-8

EPA agrees that there are uncertainties associated with each component of the sediment quality triad. For the benthic community assessment, the within-station variability in assemblages was large at both contaminated and reference stations (please refer to the response to General Issue 11.D); however, this natural variability was not so large as to obscure significant differences among locations. Although the site-specific toxicity tests indicated that responses were greatest at Stations 7 and 8, the tests also indicated adverse effects at all contaminated stations (4, 5, 7, 8, and 8A).

The evaluation of the differences in results of toxicity testing vs. the benthic community metrics at Stations 7 and 8 should include the differences in PCB exposures observed by appropriately matching effects data with exposure data. The toxicity responses at these two stations were associated with synoptic PCB concentrations (laboratory, 48-h in situ, 10-d in situ, respectively) of 213, 139, and 52.3 mg/kg tPCB (Station 7), and 72, 522, and 112 mg/kg tPCB (Station 8). In contrast, the median tPCB concentrations in the benthic community grab samples were only 3.2 mg/kg and 14.1 mg/kg, respectively. Therefore, using the most synoptic concentrations, similarly significant responses should not be expected in the benthic community metrics. Given that the concentrations in sediment from the benthic community grab samples are only modestly elevated relative to the MATC, smaller responses in benthic community metrics would be expected. This expectation is consistent with the results of the benthic community data analyses: no significant impairment in overall abundance and richness was observed at Stations 7 and 8, but there was some alteration in assemblages of sensitive taxa (i.e., chironomids).

The risk of harm is not localized to Stations 7 and 8. A comparison of sediment tPCB concentrations to the site-specific MATC value indicated that the majority of sediment in the PSA poses a risk due to PCBs. This will be presented in the revised ERA.

EPA believes that there is an acceptable dose-response relationship in most of the site-specific toxicity tests. Using the rules for determination of acceptable concentration-response described in the response to General Issue 7.D, concentration-response relationships were observed for both laboratory test species (*Hyalella* survival, growth, and reproduction; *Chironomus* survival and growth). Dose-response for the survival endpoint was also observed in 10-day in situ tests using these species, and also for a 48-h exposure using *Daphnia magna*. Consistent dose-response was not observed for 48-h in situ exposures to *Lumbriculus*, *Chironomus*, or *Hyalella*, but these results are not surprising, considering the short test duration and that the *Lumbriculus* test was intended to measure bioaccumulation and that no toxicity effect was anticipated.

EPA agrees that the TIE results are not absolutely conclusive, because such a result is not the objective of a Phase 1 TIE. However, the conclusion derived from the Phase 1 TIE (that non-polar organics are believed to be a primary causative agent in the tests) is consistent with the conclusions of the toxicity tests and other lines of evidence.

Timothy Thompson:

Based upon the presentation in the draft ERA, the conclusions are not supported. To re-summarize, (1) the tests were not designed to assess overall benthic infaunal community health in the PSA – rather they were designed to identify dose/response relationships,

RESPONSE 3.1-TT-23

EPA agrees that the primary role of the sediment toxicity tests was to enable the establishment of dose-response relationships that could be extrapolated to the remaining portions of the PSA. This approach is common in sediment risk assessments, particularly those of such a large scale, because it is not practical to conduct biological evaluations at every small area of the PSA. However, provided that the testing spans the range of concentrations and substrate types present in the PSA, it is reasonable to apply dose-response relationships drawn from a limited number of stations to the PSA as a whole. The robust sediment chemistry data set for the PSA provides a strong basis for extrapolating toxicological responses to the remaining areas of the PSA.

(2) dose-response relationships were not evident in a “large number of endpoints” – as is claimed in the ERA,

RESPONSE 3.1-TT-24

EPA believes that acceptable dose-response relationships were observed in the majority of toxicological endpoints evaluated, and will be further discussed in the process of utilizing the decision rules in the revised ERA. See also Response 3.1-RS-8 above, and the responses to General Issues 11.B, and 11.D.

(3) variability in the tPCB data need to be explored to determine if indeed the microscale differences observed are statistically different between the test stations,

RESPONSE 3.1-TT-25

Statistical comparisons cannot be conducted between stations in the toxicity studies because there is only one sediment chemistry sample per station, if only the "most synoptic data" are used (as recommended by the Reviewers). For benthic community data, there are 12 field replicates per station, such that statistical tests (including power analyses) can be performed. These analyses will be included in the revised ERA. EPA does not agree that the assessment of risk depends on specific pairwise comparisons in PCB chemistry among sites. The concentration-response evaluations and determination of an MATC were based on curve-fitting methods that do not require that each concentration can be demonstrated to be significantly different from others. Please see the response to General Issue 6.

(4) the application of median calculations to the benthic population data or the bioassay data are not appropriate,

RESPONSE 3.1-TT-26

The most synoptic concentrations were also evaluated and the results presented in Attachment D.5. Please refer to the response to General Issue 6.C.

(5) tissue residue data may not necessarily reflect sediment exposure, and that

RESPONSE 3.1-TT-27

The invertebrate tissue chemistry data were not collected perfectly synoptic with sediment concentration data. Samples for tissue analysis were collected in the general vicinity of each grab sampling location; the difference in distance between the samples was approximately 3 to 6 meters in the upstream locations and 10 to 20 meters in the downstream locations. Given the small-scale heterogeneity of PCB concentrations, these differences are sufficiently large to preclude a quantitative evaluation of tissue versus sediment concentrations using paired samples. The tissue data, however, do reflect sediment exposure as the D-net sampler was deployed in a way that sampled approximately the same sediment strata collected for the benthic community and toxicity tests. In addition, the bioaccumulation tests conducted by Wright State University indicated that tissue concentration data are related to sediment exposures, and exhibit partitioning consistent with the literature on PCB bioaccumulation.

Please refer also to Response 3.1-BS-4 and the response to General Issue 11.F.

(6) the MATC is not appropriately estimated, which effects the risk estimates for the rest of the Housatonic River.

RESPONSE 3.1-TT-28

The MATCs will be recalculated in the revised ERA following the guidance provided by the Reviewers. Please refer to the response to General Issue 6.

There was general concern noted by the Panel I concur with regarding the general risk terminology used for HQs; particularly a lack of definitions of low, moderate and high risk. HQs should be used as rough estimates of relative risk within assessment endpoints. Probability, or frequency of exceedance distributions, with defined boundaries for "low", "moderate" and "high" would be helpful in the final document. This needs to be expressed not only for tPCBs, but also for the other COCs.

RESPONSE 3.1-TT-29

EPA agrees that this discussion can be improved and will do so in the revised ERA. Please refer to the response to General Issue 7.

3.1(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made.

Valery Forbes:

The uncertainties in linking sediment chemistry to toxicity and community structure were largely addressed by analyzing different subsets of the available data (e.g., most synoptic, median). This was a useful approach. However I found very confusing the presentation of the sediment chemistry data for the toxicity and community structure samples plotted by station as it required careful reading (and explanation by EPA) to clarify that these chemical concentrations were not necessarily representative of the stations.

RESPONSE 3.1-VF-14

This discussion will be clarified in the revised ERA. Please refer to the response to General Issue 11.C.

It should be emphasized here that a substantial fraction of the 'uncertainty' is actually true variability in exposure of benthic receptor species. Such variability cannot be reduced by further measurements and should be interpreted differently in assessing risk than uncertainty due to lack of knowledge.

RESPONSE 3.1-VF-15

EPA agrees with the Reviewer's comment, and recognizes that the uncertainty associated with the benthic endpoints consists of both natural variability and incertitude. The variability in exposure is important in two main respects. First, the risks to invertebrates over a given river reach are not constant, but vary as a function of the variable PCB exposure concentrations as well as other modifying factors such as grain size or water quality. Second, the variability in exposure introduces "noise" into concentration-response evaluations, such that identification of continuous responses in individual tests (i.e., monotonically increasing effects) is not expected. Although EPA recognizes that both sources of variability exist, it was not possible to quantitatively separate these in this study. EPA will provide further clarification of this topic in the revised ERA.

Thomas W. La Point:

The only suggestion I would make on the analysis of toxicity endpoints here (see Figures D.3-12 through D.3-17) would be to use an approach increasingly used in comparative toxicity (SSDs, in Posthuma, *et al*., 2002) in which the cumulative percentage of endpoints responding are plotted against their respective concentrations inducing the response. As an example, one could re-cast Figure D.3-12 as "Cumulative percentage of endpoints responding to a given total PCB concentration." It presents a more quantitative approach, rather than having the "overview" of drawing a conclusion from trends. In the SSD approach, one could categorically state (for Figure D.3-12), "75 percent [my example estimate; not a calculation!] of the endpoints responded at a tPCB concentration of 10 mg/kg or less" and "30% of the endpoints responded at 1 mg/kg or less." The value of this suggested "cumulative response" approach is in having a quantitative statement of what percentage of endpoints are expected to respond at a given PCB concentration. As the results are now presented, there is too much latitude (in my opinion) in how one could assess endpoint response.

RESPONSE 3.1-TL-4

EPA agrees with the Reviewer's comment and will address the comments in the revised ERA by considering SSDs, multiple effect sizes, and inclusion of both acute- and chronic-based effects thresholds. Please refer to the response to General Issue 1.C.

Chronic endpoints should be separated from acute endpoints and given more weight. This should hold for the fish, amphibian, mammal and avian assessments, as well. To best understand the nature of the benthic responses, the analyses should use the synoptic chemistry data, as much as is possible.

RESPONSE 3.1-TL-5

EPA agrees with separating the acute and chronic endpoints and will also provide more weight to the analysis using the synoptic data. Please refer to the responses to General Issues 6.A and 6.C.

James T. Oris:

Uncertainties were identified and discussed within reason. Improvements can be made by discussing the uncertainties associated with using median HQ values as deterministic risk values, and discussing the uncertainties associated with multiple toxins present at levels that exceed benchmark values.

RESPONSE 3.1-JO-10

EPA will use the "most synoptic" concentrations in calculating the deterministic risk values, as recommended by the Reviewers, and will retain the discussion of the median concentrations in the uncertainty analysis.

Although EPA believes that HQs for other COCs were appropriately considered in the ERA, for several reasons discussed previously, the discussion regarding the uncertainties associated with the presence of other COCs at the concentrations observed will be expanded in the revised ERA.

Mary Ann Ottinger:

Yes, there were significant uncertainties based on sediment differences and consequential species variation. With reanalysis of the data as discussed above, this will remove some of the uncertainties and allow more straight forward interpretations.

There was a detailed discussion of uncertainties with each category within the ERA. These discussions were well done and thorough. As such, they emphasize many of the same issues that have been highlighted in my comments to this point.

Brad Sample:

Although the uncertainties are generally well captured in the ERA report, the relative importance of various uncertainties is not discussed. Interpretation and understanding of the results would benefit by such an analysis. Spatial heterogeneity of contamination is a very significant issue, that while mentioned, is not adequately discussed. Additional discussion of the potential interaction between the various contaminants in sediment on benthic invertebrates is needed.

RESPONSE 3.1-BS-18

EPA will expand the discussion of uncertainty in the revised ERA and identify, to the extent possible, those uncertainties that have the greatest bearing on risk conclusions. EPA agrees that spatial heterogeneity of contamination is a significant source of uncertainty.

Ralph G. Stahl:

There appears to be limited reflection given to the potential influence of grain size and potential seasonal changes in abundance and diversity in dynamic river systems such as the Housatonic River. On the basis of results in the coarse-grained sediments, one might conclude there were significant effects on benthic invertebrates. In contrast, where the fine-grained sediments were tested, there appears to be little or no effects on benthic invertebrates despite the fact it was these areas where some of the high levels of tPCBs were found. These points should be more fully discussed in the ERA.

RESPONSE 3.1-RS-9

EPA agrees that seasonal and inter-annual variations in benthic communities were not evaluated, and will discuss these uncertainties in the revised ERA.

EPA believes that it has appropriately considered grain size distributions in the evaluation, both in the ERA and in the additional materials supplied to the Reviewers in December 2003. However, based upon the comments from the Reviewers, additional statistical analyses to further explore the relationship of habitat factors to benthic invertebrate endpoints will be presented in the revised ERA.

EPA believes that the information presented in the ERA demonstrates that abundance and diversity of benthic invertebrates are significantly related to PCB concentrations, even after the effect of grain size distributions is considered. Effects were observed in both fine- and coarse-grained communities; the effects were observed to be larger in magnitude in coarse-grained sediment because the PCB concentrations measured synoptic with the coarse-grained benthic grabs were higher than those in fine-grained sediment. The evaluation of the differences in results of toxicity testing vs. the benthic community metrics at Stations 7 and 8 should include the differences in PCB exposures observed by appropriately matching effects data with exposure data.

Timothy Thompson:

Principal uncertainty is within the tPCB sediment measurements, as is described above. Statistical significance and effect size are not considered here, or generally throughout the ERA. Power analyses are especially important, and I recommend that this be addressed.

RESPONSE 3.1-TT-30

EPA agrees to include statistical significance, effect size, and power analyses to a greater degree in the revised ERA where appropriate. Please refer to the response to General Issue 1.C.

In addition, the ERA should provide a discussion of interlab and intralab variability on tPCB measurements, and the effect on uncertainty. At the least, the ERA should list which (how many) different labs provided tPCB measurements, methods for each lab, and provide an accounting for overall tPCB measurements per the discussion in Appendix D of the RFI.

RESPONSE 3.1-TT-31

EPA included an attachment (Appendix C.11) on analytical variability in the ERA. In the revised ERA, EPA will supplement the benthic invertebrate endpoint uncertainty assessment with relevant information from the Appendix C.11, and discuss how the choice of analytical procedures might affect risk conclusions.

3.1(h) Was the weight of evidence analysis appropriate under the evaluation criteria? If not, how could it be improved?

Valery Forbes:

As stated on p. 2-66, 'no matter what form the WOE takes, it should provide documentation of the thought process used when assessing potential ecological risk'. The weights are determined on the basis of 10 attributes that reflect the strength of association between assessment and measurement endpoints, data and study quality, and study design and execution. It is unclear how the total value for each measurement endpoint is achieved from the scores of the 10 individual attributes (e.g., Fig 2.9-1). According to the EPA's response to Panel Question VF16, the 10 attributes were considered of equal importance and the total endpoint values were determined using best professional judgment based upon the values assigned for each of the attributes. The ERA would be much more transparent if the best professional judgments were articulated more clearly.

RESPONSE 3.1-VF-16

EPA agrees and will revisit the WOE to improve the clarity of presentation. Please refer to the response to General Issue 8.

I cannot find a description of how the overall assessment within a measurement endpoint is determined. For example how are the symbols in the right-hand column of Table D 3.3 determined from the combinations of symbols for the different toxicity test results?

RESPONSE 3.1-VF-17

In this component of the WOE analyses for the benthic endpoint, the overall assessment was determined using an ordinal ranking scheme and the use of best professional judgment (BPJ) following the approach of Chapman et al. (2002). EPA is currently re-evaluating the use of this type of tabular presentation of data and may revise or eliminate the format in the revised ERA to improve clarity and transparency.

The inclusion of different numbers of effects endpoints for different species can potentially bias the WOE. For example if a species that is either very sensitive or very tolerant has more measurement endpoints than other species going into the analysis, this can lead to a biased assessment.

RESPONSE 3.1-VF-18

Please refer to the response to General Issue 6.B. EPA believes that removing redundant endpoints will eliminate the potential for bias in the assessment. The number of endpoints is similar for laboratory and in situ techniques. Although multiple endpoints are included for both laboratory species (mortality and sublethal responses), this is appropriate given that the sublethal responses (growth and reproduction) are ecologically relevant endpoints and that the number of in situ species tested (4) is twice that of the laboratory studies. Furthermore, the toxicity tests included organisms with a range of sensitivities.

Likewise when the data are scored for evidence of harm and magnitude, it seems illogical to have scores for magnitude in the event that evidence of harm is either ‘no’ or ‘undetermined’. In EPA’s response to this question (Question VF14), it is explained how such a combination of scores might be possible. This explanation should be included in section 2. Nevertheless, there must be some combinations that cannot logically occur. To follow the EPA’s example, if a field study could not rule out high risk, it would be illogical to conclude ‘undetermined/high’, because the risk could just as well be intermediate or low.

RESPONSE 3.1-VF-19

Please refer to the response to General Issue 8.A.

Thomas W. La Point:

The weight of evidence approach for these data was appropriate.

James T. Oris:

The weight of evidence analysis is not very transparent in the document as presented. Please see comment B.2 above for details.

RESPONSE 3.1-JO-11

EPA agrees to improve the clarity of the WOE evaluations in the revised ERA. Please refer to the response to General Issue 8.A.

Mary Ann Ottinger:

General Comments: The WOE is an excellent approach and very necessary in a large, complex study such as the Housatonic Rest of River ERA in which a final overall assessment must be achieved. However, in the process of arriving at the final ERA, some pertinent information is overlooked in the rush to find a single answer. This process oversimplifies a complex and dynamic situation.

This WOE is reasonable based on the data, but limited by items discussed above. Reanalysis may alter the WOE somewhat. In any event the process used to reach the WOE must be made more transparent, including more subjective professional judgments in the process.

RESPONSE 3.1-MO-8

EPA agrees. Please refer to the response to General Issue 8.A.

Brad Sample:

The weight-of-evidence (WoE) analyses is generally appropriate but is not adequately transparent.

The WoE analyses use best professional judgment to apply final weightings for the various lines of evidence and their attributes. It is unclear how final weightings were derived. Some discussion as to how these weights were derived needs to be provided – were all attributes judged equally? If not need to know why and this needs to be explained – this will allow reviewers to judge the process.

RESPONSE 3.1-BS-19

EPA agrees. Please refer to the response to General Issue 8.A.

Ralph G. Stahl:

The weight-of-evidence (WOE) analysis was appropriate, in part, under the evaluation criteria. It is reasonable but was not objective nor consistent (throughout the ERA). However, the use of professional judgment in the final assignment of risk was not clearly or fully articulated within the ERA. Numerous panel members have commented on this point previously.

RESPONSE 3.1-RS-10

EPA believes that the WOE approach used was objective, by following the approach outlined in Menzie et al. (1996). However, EPA agrees that the consistency of the application of the method should be reviewed to confirm that it was consistently applied across endpoints, and is clearly described in all cases. Please refer to the response to General Issue 8.A.

Timothy Thompson:

No. See comments above.

RESPONSE 3.1-TT-32

Please refer to the response to General Issue 8.A.

3.1(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?

Valery Forbes:

The general approach of selecting target groups based on risks observed in the PSA and downstream occurrence of the target species in combination with mapping of threshold concentrations seems logical and cost-effective. However, there seems to be some public concerns that the CT portion of the river may not have been adequately assessed. It would seem that with relatively little effort and expense, additional sediment samples could be analyzed from CT portions of the river (as recommended by Peter DeFur) which could go a long way toward alleviating these concerns and strengthening the conclusions of the risk assessment. These could be taken as one of the 'management actions' taken on the basis of the ERA.

RESPONSE 3.1-VF-20

EPA believes that the CT reaches were appropriately characterized for the risk assessment. Please refer to the response to General Issue 5.

Thomas W. La Point:

Yes.

James T. Oris:

Given the data at hand and the description of methods, risk estimates were appropriately derived for reaches of the river where site-specific studies were not conducted.

Mary Ann Ottinger:

General Comments: Studies were conducted to estimate the impact of tPCBs below PSA to 128 miles downstream (using specific endpoints and measures) [Section I.2-5]. Surveys of tPCBs indicate that there is potential for exposure risk downstream, which diminishes outside the limits of the PSA. This suggests that the potential for exposure is relatively low. As discussed below, it may be important to consider effects from mixtures of compounds and their interactions with the low level tPCBs in these regions of the river. The distribution of low levels of tPCBs is likely to be patchy, as the contaminants may be trapped in contaminated sediment.

RESPONSE 3.1-MO-9

EPA agrees that the extrapolation of the risk conclusions using MATC values requires that concentrations of other contaminants are only of value in assessing the risks for contaminants similar to those present in the test sediment. Industrial and agricultural activity is more common in Reach 7 and Reach 9 relative to the PSA; therefore, the potential for loadings of other COCs is greater than in the PSA (where the risk potential for these contaminants was found to be negligible). However, the objective of this risk assessment is to estimate the incremental risks attributable to releases from the GE facility. In the revised ERA, EPA will provide a discussion of uncertainty regarding the extrapolation to downstream areas.

These risk estimates are in line with the selected measurements of sediment and based on the samples collected considering the distance from the PSA. Conversely, if there is more agricultural activity in these areas, mixed exposure to PCBs and to pesticides may complicate the level of risk to the benthic invertebrate community. There is also uncertainty due to the variation in sediment across the habitats, with movement downstream.

RESPONSE 3.1-MO-10

Please refer to Responses 3.1-MO-1 and 3.1-MO-9 above.

Brad Sample:

Risk estimates for downstream area are based primarily on the quality and defensibility of the MATCs and their appropriate application. Comments on MATC derivation and application from 1.f apply here too.

RESPONSE 3.1-BS-20

EPA agrees and will use the revised MATCs in the descriptions of downstream risks in the revised ERA. Please refer to the response to General Issue 6.

Ralph G. Stahl:

The risk estimates for reaches of the river where site-specific studies were not conducted were objective. They may not be appropriate however. Because the MATC of 3 mg/kg tPCB is likely overly conservative, and to the extent that extrapolations of risk downstream are based on this MATC, the estimates may also be overly conservative.

RESPONSE 3.1-RS-11

EPA agrees and will use the revised MATCs in the descriptions of downstream risks in the revised ERA. Please refer to the response to General Issue 6.

Timothy Thompson:

The methodology for estimating risk in the lower Reaches is not inappropriate – I would submit that incorrect derivation of the MATC requires a reassessment of risks to the lower Housatonic. In particular, combining acute and chronic data to form a single MATC is not appropriate; these should be separate, and that was a general Panel consensus.

The Panel developed a series of "rules" that should be applied to the calculation of all MATCs for all Assessment Endpoints in this ERA. These include:

- Estimate separate acute and chronic MATCs,
- Multiple endpoints for the same receptor are not appropriate (e.g., dry-weight vs. ash-free dry weight)
- Only use the lab-measured "most synoptic data" for the bioassays
- Only those tests that display dose-response should be used in the MATC
- Use only those receptors relevant to the sediment (i.e., do not include Daphnia)
- Use all available data – not just the 6 lowest values

If MATC is equal to or lower than the background/reference concentration, then use the background number.

RESPONSE 3.1-TT-33

EPA agrees and will use the revised MATCs in the descriptions of downstream risks in the revised ERA. Please refer to the response to General Issue 6. Each "rule" described above is discussed separately in the response.

3.1(j) In the Panel members' opinions, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?

Valery Forbes:

The ERA concluded that risk is high for benthic invertebrates and that confidence in this conclusion is also high. In my view the benthic invertebrate data are more equivocal than indicated in the ERA. This is largely due to the substantial spatial and temporal variability in sediment PCB concentrations and the rather surprising (to me) difference in the relationship of taxonomic diversity versus PCB concentration between coarse and fine -grained sediments. The potential contribution of other COCs needs further attention (check especially for consistency in interpretation of HQs). One approach could be to do a multivariate analysis including other COCs. A re-analysis of the community structure data is warranted. HQs could be re-assessed as frequency exceeding the threshold. Dose-response relationships of toxicity data using most synoptic chemistry data need checking. In addition, consideration should be given to including dragonfly data, crayfish data and any other relevant data from the risk characterization that have not been included.

RESPONSE 3.1-VF-21

Please refer to the responses to General Issues 11.A (Additional Analyses), 7 (Definition of HQ and Risk Categories), 6.C (Use of Most Synoptic Data), and 4 (Use of Information from Field Studies). The risk conclusions will be modified in the revised ERA, based on the reanalysis of the MATC and the incorporation of additional analyses recommended by the Reviewers.

Thomas W. La Point:

Yes.

James T. Oris:

In the Panel members' opinion, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors? The conclusion of high risk due to PCBs to invertebrates in the PSA is supported. I believe that the other COCs warrant further attention.

RESPONSE 3.1-JO-12

EPA believes that HQs for other COCs were appropriately considered in the ERA, for several reasons:

- The HQs for other COPCs were based on conservative screening SQVs, which have been demonstrated to yield a high percentage of false positives (Chapman and Mann 1999). This conclusion is supported by the ERA findings, which indicated a site-specific tPCB MATC for benthic invertebrates that is more than an order of magnitude higher than the median literature-based SQV. Therefore, the caveats attached to the use of SQV-based HQ screening values, for both PCBs and other COCs, are appropriate and consistent with the sediment quality triad approach.

- The HQs were not used to make definitive conclusions regarding risk, but instead were used to indicate order-of-magnitude differences across contaminant types. In this context, the HQ values for other COCs were compared to PCBs, and the differences were observed to be large (i.e., two orders of magnitude). EPA believes that such large differences are sufficient to rank PCBs higher in terms of potential toxicity, and agree that the HQs cannot be used to make definitive statements of absolute risk.

- Other COCs with HQ>1 were not eliminated from further consideration based on the HQ results. Other COCs were retained in the ERA, and other lines of evidence, such as TIE results and correlations of chemistry to effects measures, were used to make risk conclusions.

Also please see the response to General Issue 7 that describes the different uses of the term “HQ” in the ERA.

Mary Ann Ottinger:

General Comments: Based on the information provided in the ERA, the conclusions about risk for representative species are premature and require some additional information. This is due to the potential underestimated risk associated with the local sediment environments and particle size of the sediment that would result in varying concentrations of tPCBs.

RESPONSE 3.1-MO-11

Because EPA concluded that risk to benthic invertebrates was high, this conclusion would remain the same even if the ERA were made more conservative (i.e., protective), as the Reviewer recommends.

Brad Sample:

In my opinion the conclusions drawn from this section are probably correct. Definitive conclusions cannot be made however until reworking of the data, as suggested above, is complete.

RESPONSE 3.1-BS-21

EPA will revise and expand the data analysis supporting the benthic invertebrate risk assessment in the revised ERA. Please refer to the responses to General Issues 6 and 11 for a summary of the major areas of revision.

Ralph G. Stahl:

The evaluation supports the conclusions regarding high risk to local populations of benthic invertebrates; however, this risk is localized to specific stations within the PSA. The estimates of exposure, based on the data selected for this purpose, and the method used to derive the MATC, are overly conservative. This will drive the HQs higher and may indicate a higher level of risk than is present.

RESPONSE 3.1-RS-12

EPA believes that the risks to benthic invertebrates demonstrated at specific locations can be appropriately extrapolated to other locations in the study area with measured tPCB concentrations, and by extension, to contiguous sediment environments with similar habitat characteristics. Please refer to Response 3.1-RS-8 above. EPA agrees that recalculation of the MATC is appropriate (please refer to the response to General Issue 6).

In my opinion, there is clear evidence of a risk of toxicity to benthic invertebrates residing at stations 7 and 8 in the PSA, yet this is not fully reflected in the abundance and diversity observations from the field. The conflict between toxicological results, abundance of benthic invertebrates, and sediment chemistry results is not unique to the Housatonic River. It is also evident that, in the final evaluation of risk, sediment chemistry (tPCB content in bulk sediment) was given significant weight compared to the benthic diversity and abundance results.

RESPONSE 3.1-RS-13

EPA agrees that it is often the case that the elements of the SQT can result in conflicting lines of evidence. However, EPA believes that the benthic community assessment and toxicity data are reasonably consistent, when the synoptic PCB concentrations in each study are considered. A clearer discussion of the linkage between components of the triad will be provided in the revised ERA. Please refer to the response to General Issue 8.A and Response 3.1-RS-8 above.

EPA believes that the sediment chemistry data were given appropriate weight compared to benthic diversity and abundance results. Sediment chemistry was afforded the lowest weighting of these three triad components. The WOE weighting table (Table 3.4-1) shows that the overall weighting for sediment chemistry was "low/moderate," whereas the site-specific studies ranged in weighting from "moderate" to "moderate/high." Chemistry data were used primarily as indicators of exposure and for concentration-response evaluation and evaluation of relative contributions of COCs. The ERA specifically emphasized the limitations to the application of sediment quality criteria for both PCBs and other COCs.

Timothy Thompson:

No, a reanalysis of the data is required before final conclusions can be made of the infaunal data. These include the spatial variability of tPCBs at each station, pairing the appropriate tPCB measurement with measurement endpoint, evaluation of substrate type on benthic infaunal communities, appropriate indicies for infaunal communities, dominant taxa in the community structure, frequency distributions of tPCB and other COCs relative to TRVs and/or MATCs, and a better explanation of professional judgment used to assigning the weight-of-evidence factors.

RESPONSE 3.1-TT-34

As described in the responses above, EPA will reanalyze the data in response to the Reviewers' comments. Please refer to the responses to General Issue 11 (spatial variability of PCBs, substrate type, appropriate indices, additional analyses), 6.C (Use of Most Synoptic Data), 1.C (Statistical Methodology), and 8.A (Weight-of-Evidence Approach).

3.2 Amphibians

3.2.(a) Were the EPA studies and analyses performed (e.g., field studies, site-specific toxicity studies, comparison of exposure and effects) appropriate under the evaluation criteria, and based on accepted scientific practices?

Valery Forbes:

Generally yes. In principle I believe it could be efficient to use some of the field studies performed for site characterization in the risk assessment (e.g., vernal pool surveys for breeding amphibians Appendix A.1). Unfortunately these were concluded to be an insensitive tool for detecting effects of PCBs.

RESPONSE 3.2-VF-1

Risks to amphibians in the PSA and other parts of the Housatonic River floodplain were characterized using data from a variety of sources, ranging from site-specific toxicity tests to literature reviews of amphibian toxicity research. Each source of data provided information useful for characterizing risk, although the relative value of the data varied. Surveys of temporary (vernal) and permanent pools in the PSA were performed in 1998 to identify species presence, breeding activity, and use by habitat. During these surveys, quantitative and qualitative data were collected using standardized methods and data forms; however, most sites were only visited once. To minimize potential sampling bias, only data collected from the first site visit were used from pools with multiple visits. This allowed comparison of species richness across pools. The first visit was conducted when pools were characterized using the Massachusetts Natural Heritage and Endangered Species Program (MNHESP) methods. These data were considered to be an insensitive tool for detecting effects of PCBs for several reasons:

- Species behavior. Richness data for many sites were collected only during one sampling event. Secretive species, or those that are active during periods outside of the survey window, were less likely to be observed.

- Spatial heterogeneity of contamination. As illustrated in Figures 2.5-5 to 2.5-11, contamination concentrations in the floodplain vary on a scale smaller than the home ranges for many amphibians in the floodplain. This means that exposure varies greatly depending on the individual, species, and location in the floodplain, and this may result in a high degree of variability that may not be sensitive to detection using the MNHESP survey methods.

- Presence of uncontaminated amphibian habitat adjacent to contaminated sites. Suitable amphibian habitat occurs outside of, but immediately adjacent to, the floodplain. These areas are within the home range of amphibians, providing opportunities for movement in and out of the PSA.

As stated above, I believe that the definition of assessment endpoints for amphibians is inappropriate.

RESPONSE 3.2-VF-2

Please refer to the response to General Issue 2.

The design of both the leopard frog and wood frog site-specific toxicity tests (FEL 2002) was rather involved and therefore somewhat difficult to follow. In both studies an excellent gradient of sediment PCB concentrations in the test pools was achieved. However, it was determined that exposure of egg masses and young was largely via maternal transfer and not pool sediment which, to some extent, complicates interpretation of the early life stage results.

RESPONSE 3.2-VF-3

EPA agrees that interpretation of early life stage effects needs to be carefully evaluated. Fort Environmental Laboratories (FEL) (2002a) found no relationship between wood frog egg mass PCB concentration and egg mass characteristics such as weights, egg counts, egg necrosis, fertilization rates, and hatching success. In leopard frogs, however, FEL (2002b) observed that almost all of the frogs collected from the PSA could not be fertilized in the lab because oocytes were not fully developed.

In the site-specific toxicity study of leopard frog reproductive success, it was a weakness that no frogs were captured from the reference area and that the study had to rely on purchased frogs for the control group. Thus, the reference group is not a true control and should be dropped from the statistical comparisons. In this same study there were found low stage VI oocytes at all stations which was suggested could be due to frogs moving among sites (questioning actual exposure-response relationships). There was also a very small sample size available with only one to a few egg masses collected per pond.

RESPONSE 3.2-VF-4

EPA did not perform inferential statistical analysis for concentration-response relationships for the leopard frog effects data (Appendix E, Section 3.2.1). Hypothesis tests were, however, performed for cross-over and spiking studies, which used larvae hatched from leopard frogs purchased from Carolina Biological Supply (CBS). EPA believes that this was appropriate because the study was able to compare the effects of exposure to reference site water and sediment to contaminated site water and sediment from the PSA on larval leopard frog mortality, metamorphosis, and malformation. The frogs obtained from CBS were collected in Vermont and verified by FEL as northern leopard frogs. Even though these frogs were not collected from the Housatonic River watershed, EPA believes they nonetheless provide a valid comparison for expected effects to northern leopard frogs. Please also refer to the response to General Issue 12.A.

Thomas W. La Point:

Yes, there was a good attempt to gather information on the exposure of the two ranid frog species. The laboratory toxicity studies, with associated larval endpoints, provide particularly strong evidence. The EPA leopard frog field study unfortunately led to small sample sizes (this may be indicative of long-term population trends in the Housatonic River Valley). However, as per page 4-32 of the 4.4.1.1.1 section, using "visual interpretation" to understand results, particularly with small data sets, can lead to dramatic errors in interpretation – if only because of the sampling variance inherent with small sample sizes.

RESPONSE 3.2-TL-1

EPA carefully considered sample size, visual interpretation of data, variability of data, and the uncertainty associated with findings when using these data to evaluate risk to amphibians. Appendix E, Section 3.2.1, details the process used for evaluating leopard frog effects data. Inferential statistics for concentration-response relationships were not considered appropriate for the analysis of leopard frog effects data for the following reasons:

- The magnitude and range of responses observed were not well suited to the detection of a concentration-response relationship; the leopard frogs tended to exhibit much more of a threshold response for various endpoints. These types of data distributions are better evaluated via visual interpretation (figures) and examination of average response data (for a given endpoint) at each station.

- Sample sizes were small, primarily as a result of the low field availability of test animals. A sample size of 5 was common with these data, and although correlation analyses can be conducted, the relationship must be very strong ($r_{s,0.05,3}$ = 0.878) to be considered statistically significant. This is an unrealistic expectation for biological response data.

- The number of meaningful statistical tests that could be conducted was small, primarily due to the necessity of revising the study design. Because fertilization of the field-collected females was unsuccessful, there was no biological relationship between female PCB body burdens or reproductive tissue endpoints (i.e., percent Stage VI oocytes, percent malformed sperm cells) and larval/developmental effects endpoints (i.e., incidence of metamorphosis).

James T. Oris:

Opinion is generally yes. More effort seems to have been exerted on amphibian individual-level and population-level effects than some of the other receptors, and it is not clear why. Why were deformities and sex ratios considered important for amphibians, but not other species? Why were development and maturation used for amphibians, but not other species? Also note, the statement of assessment endpoint in highlighted box (p. 4-1) does not match the formal statement of the assessment endpoint.

RESPONSE 3.2-JO-1

Deformities and sex ratios were selected as measurement endpoints by FEL during the study design because of their previous experience with amphibian toxicity research. This experience indicated sensitivity to these endpoints, particularly for COCs believed to be endocrine disrupters. EPA will include a more detailed discussion of measurement endpoint selection in the revised ERA, and will revise the text in the highlighted box on page 4-1.

Mary Ann Ottinger:

There were several studies, both in the field and in the lab. In total, these studies provide useful information for the ERA and lead to reasonable risk assessments for these populations. However, there are flaws and limitations in many of the studies, which make them difficult to interpret and somewhat confusing. The leopard frog study is limited by small sample size and by lack of reference site animals.

RESPONSE 3.2-MO-1

EPA agrees that the leopard frog study results have some limitations, which were carefully considered during the preparation of the ERA. Appendix E, Section 5.1, discusses uncertainty associated with the leopard frog study. Please also refer to the response to General Issue 12.A.

This results in heavier reliance on evidence from the wood frog. The wood frog appears to be an appropriate species because of its use of vernal pools containing sediment; however the time of PCB exposure is limited to larval stages; therefore the studies were conducted [I.4-15 (Figure 4.2-3] to assess the COC impact on eggs in the vernal pools. The amphibian community study supports the impact of tPCBs on species richness, with wood frog larval stages most impacted. Finally, if there was a high rate of deformities in the tadpoles, these individuals would not be likely to show up in the adult population, thereby underestimating the rate of deformities (DELTs). In line with this comment, there were no observations on failed metamorphosis, which would indicate thyroid system abnormalities, potentially due to the PCBs.

RESPONSE 3.2-MO-2

The wood frog study examined the effects of PCB exposure to egg masses, early and late larval stages, and metamorphosis. The potential impacts of COCs were evaluated for multiple endpoints for each (e.g., rates of fertilization, rates of hatching, larval growth and development) of these stages of development. EPA concurs with the conclusion by Dr. Fort and the Reviewer that it is likely that deformed larvae would not survive to maturity, and therefore would not be observed in the community survey, thereby underestimating the rate of DELTs. In the laboratory study, tadpoles that failed to metamorphose completely and died were counted as dead larvae. When dead metamorphs were encountered in the field, they were not collected for shipment because the sampling protocol required the collection of live individuals. Also, it is difficult to observe dead metamorphs in the field because of their dull coloration, small size, and the fact that they rapidly decompose.

Brad Sample:

The field and laboratory studies conducted by EPA were suitable and appropriate and based on accepted scientific practices. The laboratory-based leopard frog study was only partially successful – laboratory fertilization of eggmasses derived from field collected frogs did not work. This reduces the strength of the data. Additionally, the use of reference frogs from VT adds some uncertainty, but because these frogs were collected from within the region and same time of year, I believe that they are suitable and appropriate.

Response 3.2-BS-1

EPA concurs that the use of the control frogs collected in Vermont was suitable and appropriate. Please also refer to the response to General Issue 12.A.

Ralph G. Stahl:

Field studies conducted by EPA were appropriate under the evaluation criteria. The inability to find adult leopard frogs in the reference area for use in the FEL-conducted studies is puzzling. There is no clear explanation for the absence of adult leopard frogs in the reference area during the collection period. The collection may have been conducted inappropriately but there is no evidence that this was problematic during previous field work. Unfortunately the inability to find adult frogs casts substantial doubt on the usefulness of the FEL leopard frog toxicity studies, as noted below.

RESPONSE 3.2-RS-1

EPA carefully considered the limitations and uncertainty associated with some of the leopard frog data and, while the limitations are acknowledged, believes that the data add to the weight-of-evidence regarding risk to amphibians. Please also refer to the response to General Issue 12.A.

To compensate for the absence of adult leopard frogs in the reference area, FEL purchased adult frogs from a commercial supplier. As a result, these adult frogs were used in lieu of true "reference area" frogs. One conclusion drawn from the FEL toxicity studies was that the high predominance of immature oocytes in female leopard frogs resulted from their exposure to tPCBs and TEQ in the PSA. However, information supplied by GE during the January 2004 public meetings suggest that EPA's finding of a predominance of immature oocytes in leopard frog females from the PSA may be due to low temperatures at the time of collection rather than as a result of exposure to tPCBs. In effect the female frogs in the PSA appear to have been collected at less than optimal temperatures. This suggests that the leopard frog field collection may have been conducted at a time when oocytes had not yet developed more fully.

RESPONSE TO 3.2-RS-2

EPA does not agree with GE's assertion that leopard frogs were collected at the wrong time of year. The timing of the collection of adult leopard frogs from the target stations coincided with the normal onset of reproductive receptiveness and initiation of breeding activity. Adult specimens were collected between March 25, 2000 and April 22, 2000. Surface water temperatures in the PSA were approximately 8 to 10°C at this time. These temperatures represent the ideal environmental "triggers" for the frogs to emerge in the early spring and gather in breeding areas. Typically, males begin chorusing when water temperatures reach approximately 8°C, with oviposition peaking when water temperatures reach 10°C (Gilbert et al. 1994). Hine et al. (1981) reported the occurrence of breeding when water temperatures reached or slightly exceeded 10°C in Wisconsin ponds.

From April 21 to 25, 2003, GE contractors observed leopard frog egg masses in pools within the PSA. This was the first week of their leopard frog surveys, which coincides with the same time period in which EPA contractors observed leopard frog egg-laying. Based on these observations and the similar timing of the 2000 surveys, it is unlikely that the female leopard frogs were in the early stages of oocyte development or that they had already bred. To the contrary, it is likely that the 2000 surveys were conducted during the optimal breeding period.

EPA will add text to the revised ERA to clarify the appropriateness of the timing of leopard frog collection. Please also refer to the response to General Issue 12.A.

Therefore, site-specific leopard frog toxicity studies conducted by FEL were not based on accepted scientific practices. For example, adult frogs were not obtained from the reference area. Instead, adult frogs were purchased from a commercial supplier and utilized in the laboratory testing. Although this deviation from standard practice was detailed in the ERA, the results of this specific study are, in my estimation, only qualitative in nature. Direct comparisons between frogs obtained from the PSA and the reference area cannot be made since the commercially purchased frogs (in lieu of reference area frogs) were not exposed to the site-specific environmental conditions that would have been present in the reference area.

RESPONSE 3.2-RS-3

In the absence of frogs collected from reference areas, external control frogs of the same species collected a few hundred miles away from the PSA provide the most relevant animal for analysis of “clean” media in cross-over and spiking studies. The use of purchased control frogs is more relevant than data from the literature because the purchased frogs were subsequently exposed to sediment/water from the reference site (Muddy Pond). The fertilization of the control frog egg masses and the subsequent laboratory culture of the larvae through metamorphosis in Muddy Pond media represent a valid reference condition exposure scenario. These larvae spent over 3 months (106 days) in the Muddy Pond sediment/water. Therefore, the EPA sees no reason to exclude these reference developmental data from the evaluation of concentration-response relationships. Please also refer to the response to General Issue 12.A.

The separation of the PSA into areas based on low, medium and high tPCB content in the sediments, appears to be appropriate. There appears to be a correlation, albeit weak, between the tPCB content in sediments and the whole body burden of tPCBs in frog tissues. In contrast, the FEL study showed a poor dose-response between tPCBs in sediments and the percentage of abnormal sperm in adult male leopard frogs. This lack of a dose response is also evident from the FEL toxicity study where the percentage of stage VI oocytes in adult females at all sampling locations is significantly less than the reference.

RESPONSE 3.2-RS-4

As discussed in Section E.3.2.1 of the ERA, responses observed in the leopard frog were more similar to a threshold-level response than a monotonic dose-response. In addition, sample sizes were small, due to restricted field availability of test animals. This limited the number and types of statistical tests that could be performed. Graphical comparisons were presented to illustrate potential

> relationships between effects and tissue and sediment tPCB concentrations. Although these types of comparisons are qualitative, they are useful in the overall evaluation of all of the amphibian data. The uncertainty associated with the use of these qualitative data was discussed in Section E.5.1 of the ERA.

As noted previously, the adult female leopard frogs used in this toxicity test were purchased commercially and thus makes highly uncertain any comparisons and conclusions that might be drawn from this particular study.

Timothy Thompson:

A primary initial issue to address is the assessment and measurement endpoints – as the conduct of the studies relate directly to this. A general issue relates to use of maturation and development as an assessment endpoint. The Panel wanted to understand why "maturation" and "development" were important (as opposed to other indicators such as survival, sustainable populations, etc), and when queried, EPA responded that it was what the parties[1] agreed to. Several Panel colleagues were eloquent and passionate about "maturation" not being an assessment endpoint – arguing that if protect of populations of amphibians is the goal, then the assessment and measurement endpoints need to be appropriate parameters relating to the population (or societal values). Impairments can occur related to maturation, but that may not effect populations. I agree with their general comment, but believe that if the biological link is made between maturation and population – then that condition is effectively addressed. Recommend that the final ERA work to make that connection.

> **RESPONSE 3.2-TT-1**
>
> The development and specific wording of the Assessment Endpoints has a long history involving GE and several regulatory agencies. EPA believes it is not appropriate to redefine the endpoints at this late stage of the ERA process. Please refer to the response for General Issue 2 for additional detail.

In general, the studies were well thought out, and executed. Unfortunately, not all of the studies proceeded as planned and I felt the study authors made the best use of data possible, even while not always agreeing with the conclusions. For example, a specific weakness cited in the Leopard frog reproductive success cited by GE and some of my Panel colleagues was that no frogs were captured from the reference area and that the study had to rely on purchased frogs for the control group. The suggestion has been made that since the reference group is not a true control it should be dropped from the statistical comparisons. I disagree with this position. I do not agree with this position; there is meaningful information to be derived from using the control frogs from outside the PSA. It adds a higher degree of uncertainty within the final conclusions, but recommend that those be retained in subsequent statistical analyses and assessment of risk.

---

[1] Although a formal BTAG was apparently not convened, I took "the parties" to mean the other federal (NOAA, USFWS) and State (Massachusetts and Connecticut) agencies. Recommend that a memo documenting the development of the Assessment and Measurement endpoints with the appropriate resource agencies and trusties be added as a memo or appendix to the Final ERA.

RESPONSE 3.2-TT-2

EPA concurs that the use of control frogs from outside the PSA was appropriate even though there were some uncertainties associated with these data. Please also refer to the response to General Issue 12.A.

3.2(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria, based on accepted scientific practices, and incorporated, appropriately in the ERA?

Valery Forbes:

Although I am not an expert in amphibian field studies it seems that the field studies performed here (i.e., leopard frog egg mass surveys) were not particularly powerful tests of potential PCB effects on frog populations due to problems linking actual exposure to observed effects and to small sample size.

RESPONSE 3.2-VF-5

EPA agrees that the leopard frog egg mass survey was not a particularly powerful test of potential PCB effects on frog populations due to problems linking exposure to effects. In their application for the collection permit, GE indicated that the purpose of the collection was to conduct an in situ experiment on the reproduction and development of Northern Leopard frogs, with possible chemical analyses. However, this aspect of the study was not implemented. EPA believes that the egg mass survey, however, has some application as part of the WOE evaluation.

The wood frog study by Resetarits (2002) seems to have been well designed (i.e., randomized complete block design, large numbers of larvae per treatment), but did not adequately simulate exposure of frogs to PCBs in the field (i.e., which would include both maternal transfer and sediment exposure).

RESPONSE 3.2-VF-6

EPA agrees with the Reviewer's assessment. Maternal transfer was the only pathway of PCB contamination assessed in GE's study, and there are other known pathways of larval anuran PCB contamination. In particular, exposure through sediment contact throughout the entire larval period may be the most important exposure pathway affecting larval amphibian development. The GE study did not adequately address either of these factors. The problem of an inadequate contaminant exposure scenario more than outweighs the perceived benefits of an in situ experiment or explicit consideration of density dependence. This fault is magnified by the use of relatively uncontaminated pools for the deployment of study animals and the restriction of exposure of the study animals to pool sediment.

Thomas W. La Point:

The field surveys were conducted and follow accepted scientific practices. However, the duration of the GE wood frog field study was relatively short and, given the limited numbers of egg masses and limited numbers of ponds in the contaminant distribution classes. This is not, *per se*, incorrect – but the study has low power. The use of the field tests, with the associated variances in distribution and body burdens, should not be used to mask the more direct results from laboratory toxicity studies by Fort, *et alia*. Although the laboratory toxicity assessments may be deemed conservative, the potential for chronic effects in the field must be taken into account, even if the field data were "inconclusive."

RESPONSE 3.2-TL-2

EPA agrees that the GE wood frog study has low power.

James T. Oris:

Why were no reference areas used? Why did they not evaluate relationship between number of adults versus number of eggs? The deficiencies in the GE studies were appropriately noted and the studies were generously incorporated into the ERA.

RESPONSE 3.2-JO-2

EPA concurs with the Reviewer's assessment. From the description included in the application for the collection permit, it appears that GE intended to conduct additional measurements, but did not implement the study.

Mary Ann Ottinger:

The survey of leopard frog eggs was limited by selection of appropriate sites and lack of reference sites. The number of sites associated with the 4 ranges of tPCBs was small and led to conclusions of no differences associated with level of contamination. More study would be needed in order to definitively make this conclusion based on the other data collected in amphibians in the PSA.

RESPONSE 3.2-MO-3

EPA concurs with the Reviewer's assessment that more study would be needed to substantiate the conclusions made in GE's study.

The wood frog field study Resetarits (2002) was well designed, but had an additional variable of density, which was exacerbated by predation in all the sites. Differential density pressures among replicates add to the variability and make the data more difficult to interpret. In addition, this study was not included in the assessment because exposure did not occur during the late larval stage (most sensitive stage); moreover, little effect was attributed to maternal transfer. The study is difficult to interpret or even analyze due to the lack of reference sites.

RESPONSE 3.2-MO-4

EPA agrees that the GE study results were difficult to interpret, and that the lack of reference sites limits analysis of findings.

Brad Sample:

GE conducted two amphibian studies – one field survey of occurrence of leopard frog eggmasses in relation to PCB concentrations and one in-situ study of the interaction between PCBs and population density among wood frogs.

Both GE amphibian studies suffered from several statistical and design issues – selection of treatment groups following data collection (leopard frog study), small sample sizes (both studies), no references or controls (wood frog study), etc. As a consequence, their utility in addressing the risk-related questions was reduced. Given these issues, I feel that data from these studies was used and incorporated into the ERA appropriately.

Ralph G. Stahl:

The GE studies conducted on the leopard frog were appropriate under the evaluation criteria and were incorporated appropriately in the ERA. The leopard frog egg mass field survey conducted by GE failed to quantitatively evaluate the number of fertilized eggs within the egg masses; however, this oversight was discussed in the ERA and the conclusions that could be drawn from this qualitative evaluation were appropriate.

The wood frog study conducted on behalf of GE by Dr. Resetarits was appropriate under the evaluation criteria and was incorporated appropriately in the ERA. The conclusions drawn from this study suggest that wood frog populations are density dependent which does not seem to be accounted for properly in the modeling projections. Because the results of the wood frog studies by FEL and GE are the most comprehensive and potentially the primary basis for estimating potential risks to amphibians, I recommend that EPA re-evaluate the Resetarits study as well as the modeling projections conducted for EPA.

RESPONSE 3.2-RS-5

EPA gave careful consideration to the findings in the Resetarits report during the preparation of the ERA; however, there were a number of shortcomings that limited the use of the data, as explained in detail in Response O-RS-21. Limitations included lack of detailed description of methods, evaluation of maternal transfer only, and premature conclusion of the study. Please refer also to the response to General Issue 12.B.

Timothy Thompson:

GE's studies are well thought out and executed. The unfortunate aspect of their work is that it appears to have been conducted outside of the overall ERA Problem Formulation process, and more as a response (counterpoint) to what the ERA was testing. For example, the Leopard Frog study appears to have been appropriately designed to simply answer the question "are leopard frogs in the PSA and are egg masses present" -- as opposed to statistically powerful tests of whether leopard frogs are actively reproducing within the PSA. Notwithstanding the very sample size, and the accompanying low statistical power, the study did demonstrate the Leopard Frogs are indeed in the PSA with associated egg masses. That is an effective counterbalance to the EPA study.

RESPONSE 3.2-TT-3

EPA agrees that leopard frogs breed and live in the PSA, which was one of the reasons that this species was chosen as a representative species for the ERA. EPA believes that the egg mass survey was not an effective counterbalance to the FEL studies for the reasons stated in Appendix E, Section 5.3, and the fact that GE found 2.4 egg masses per hectare (ha), in comparison to the 277 egg masses per ha reported by Hine et al. (1981) or the 58 egg masses per ha reported by Gilbert et al. (1994). The GE study, as implemented, was biased toward only those individuals that are reproductively fit, and made no attempt to determine if there were adult frogs that were not reproductively fit. The presence of a species in a contaminated area is not proof of absence of an effect, particularly for a species that is mobile, such as the leopard frog, and also occurs in an area with heterogeneous contamination. In the revised ERA, EPA will further discuss the limitations of the GE leopard frog study, particularly the finding of no harm.

3.2(c) Were the estimates of exposure appropriate under the evaluation criteria, and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?

Valery Forbes:

Some uncertainties in exposure in some of the field studies as indicated above. No issues with COCs.

Thomas W. La Point:

Given the ability of frogs to range easily over one or two ponds, the estimates of exposure from sediments and food organisms is appropriate. However, the lack of toxicity information on salamanders (generally thought to be more sensitive than frogs) should lead to precaution in estimating HQs too closely in these studies. The variances in effect concentrations may be quite large.

RESPONSE 3.2-TL-3

EPA agrees that some amphibian species may be more sensitive than the two frog species chosen as representative amphibians for the ERA as described in Appendix E, Section 4.8.2. Please also see the response to General Issue 7.

James T. Oris:

This section was appropriate.

Mary Ann Ottinger:

The amphibian community assessment endpoint relies primarily on field data from the wood frog and leopard frog with additional studies. Accordingly, it would have been valuable to have data on species richness in various portion of the PSA beyond the initial survey to relate to the findings from these studies.

RESPONSE 3.2-MO-5

EPA agrees that additional data on species richness beyond the initial survey would have contributed to the understanding of the potential effects of PCBs on amphibian populations. These data, however, are time-consuming to collect, and natural variability resulting from the heterogeneity of contamination, mobility of different species and individuals, and inherent stochasticity means that there will be uncertainty associated with the findings. This was one of the reasons that EPA chose to perform more site-specific toxicity tests for amphibians.

The leopard frog is a good model in terms of more potential contact with the COCs. However, several questions/issues should be addressed to minimize uncertainties in these data.

1) Were the number or species surveyed at various sites sufficient to assess differences at the PSC?

RESPONSE 3.2-MO-6

EPA agrees that the leopard frog was appropriate for estimating exposure to amphibians from COCs in the PSA. The number of sites surveyed as part of EPA's amphibian work was sufficient to assess differences in the PSA due to COC concentrations. Uncertainties arose from other factors, including the mobility of frogs between areas with varying concentrations of COCs, as described in Appendix E, Section 5.

2) Were there appropriate reference sites in the field surveys and for field samples? The laboratory tests used appropriate measurement endpoints for ascertaining PCB impact in the wood frog. However, the sperm abnormalities data were not statistically analyzed due to a small n.

RESPONSE 3.2-MO-7

There were no reference areas for GE's amphibian field surveys. EPA used reference areas for FEL's wood frog and leopard frog studies; however, as described in Section 5.1 of Appendix E, the absence of leopard frogs in the reference areas limited the number of endpoints that could be evaluated. Few statistical tests could be performed because of the limited number of endpoints evaluated. EPA believes the wood frog reference areas were appropriate.

3) There is an assumption that the histological evidence indicates that the adults would be sterile; however there was no attempt to assess these animals as adults and no endocrine parameters were measured to determine if these animals were affected. Therefore, this is not a legitimate assumption even if it appears to be correct direction! In total, the risk characterization not obvious based on the data.

RESPONSE 3.2-MO-8

In the opinion of the Principal Investigator (Dr. D. Fort), these malformations would likely result in sterility in adult frogs.

4) If salamanders may be at greater risk, why were they not studied, even in a laboratory study?

RESPONSE 3.2-MO-9

ERA endpoints, for the most part, were developed and agreed to in the discussions preceding the Consent Decree, which influenced the development of the work plan under which the ERA was conducted. At that time there were limited effects data for salamanders, and it was decided that the lack of toxicity data would hinder development of a species-specific risk assessment for salamanders as representative species. Please refer also to the response to General Issue 2.

6) Lack of a reference area control presents some uncertainty, but is the best option in the situation that occurred.

7) As in the birds, the identification of the PCB congeners should point to some specific endpoints that may be more sensitive measures of exposure and would complement other data collected in the reproductive laboratory study.

RESPONSE 3.2-MO-10

PCB congeners were analyzed in FEL's leopard and wood frog studies, as described in Appendix E, Sections 2.7.4.1 and 2.7.4.2. Although several amphibian studies have investigated the effects of PCB congener 126 on leopard frog hatching success, development, and metamorphosis (Rosenshield et al. 1999; Beatriz et al. 2000), amphibian effects data for other specific congeners are limited. A more complete description of the PCB congener data will be included in the revised ERA.

8) Finally, mixtures of chemicals are of concern for the amphibians, especially those whose habitat is in the proximity of agricultural activity.

RESPONSE 3.2-MO-11

The impacts of multiple COCs were investigated as part of the amphibian ERA as described in Section E.2.1.1. Available information on potential synergistic effects of chemical mixtures on amphibians was considered during the risk characterization. Unfortunately, because the body of literature on amphibian risk is still not large, there was not much pertinent information from sites similar to the Housatonic River PSA. Please refer also to the response to General Issue 3.

Brad Sample:

The exposure evaluation appears to be suitable and appropriate.

Ralph G. Stahl:

The estimates of exposure were appropriate under the evaluation criteria, as was the refinement of analyses for the COCs.

Timothy Thompson:

Generally, yes. Within the context cited under (a) and (b), above, the estimates of exposure are appropriate within both EPA and GE studies.

3.2(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria?

Valery Forbes:

The relationships between metamorph malformations, sex ratio and population-level effects were not quantified which makes interpretation of the seriousness of effects on the measured endpoints difficult.

RESPONSE 3.2-VF-7

Attachment E.3 in Appendix E, the wood frog population model, integrated malformation data, sex ratios, and population-level effects. Text in this section of the revised ERA will be expanded to more clearly explain the model inputs, outputs, sensitivity of certain input parameters, and potential influences of density-dependence.

Also see points on derivation of MATCs for invertebrates.

RESPONSE 3.2-VF-8

Please refer to the response to General Issue 6.

Thomas W. La Point:

Yes!

James T. Oris:

There is a need to clarify terminology for frog studies. Controls=frogs from Carolina Biological Supply. Reference=frogs from reference ponds.

RESPONSE 3.2-JO-3

EPA will clarify the terminology for frog studies in the revised ERA. Please refer also to the response to General Issue 12.A.

Discussion of choice of MATC values is warranted. A sediment MATC of 3mg/kg was set even though EC20 values were much lower than this, especially the EC20 for altered sex ratio.

Altered sex ratio should be considered a significant effect since literature studies indicate that transformed "females" are likely not fertile. Further examination of this data is needed.

RESPONSE 3.2-JO-4

The sediment MATC of 3 mg/kg tPCB used in the ERA is just below the EC20 values for Phase III malformed metamorphs (3.27 mg/kg). A moderate effect for metamorph malformation is believed to be biologically more relevant than the sex ratio EC20 because one male can potentially breed with multiple females. Ouellet (2000) also suggests that malformation rates greater than 5% are biologically relevant. In addition, 3 mg/kg tPCBs is within the 95% confidence limit of the EC20. Phase I metamorphs external (only) malformations ranged from 13 to 17%, just below the moderate effects level, but above the biologically relevant level; however, inclusion of internal malformations would have increased these numbers and allowed probit modeling. Please refer also to the response to General Issue 6.

Mary Ann Ottinger:

The wood frog studies would have been more complete if thyroid hormones had been analyzed as a specific responsive hormone to PCBs.

RESPONSE 3.2-MO-12

EPA agrees that thyroid hormone analysis and other analyses may have provided additional useful information on potential toxicological effects that could have been incorporated into the ERA. Practical limitations necessitated that EPA and GE develop study plans for representative endpoints using methods believed to be suitable for characterizing risks.

Also, the number of tadpoles that did not metamorphose properly or were abnormally metamorphosed would have provided additional information.

RESPONSE 3.2-MO-13

In the laboratory study, tadpoles that failed to metamorphose completely and then died were counted as dead larvae. When dead metamorphs were encountered in the field they were not collected for shipment because the sampling protocol required the collection of live individuals. Also, dead metamorphs were very difficult to observe in the field because of their small size, dull coloration, and the fact that they rapidly decomposed.

If possible, the laboratory studies would have been stronger if a dose-response relationship had been examined, especially relative to selected measurement end points that are reliable indices of PCB exposure in amphibians.

RESPONSE 3.2-MO-14

Section 4.3 in Appendix E, Concentration-Response Analysis – Toxicity Endpoints, contains a detailed description of the dose-response relationships examined, describes the process of endpoint screening, and describes the methods used and the statistically significant results for dose-response relationships in wood frogs.

Brad Sample:

The effects data appear to be suitable and appropriate.

Ralph G. Stahl:

The effects metrics were appropriate under the evaluation criteria. The large number (up to 11) of effects measurements provides substantial coverage of potentially problematic physiological and morphological impacts. Impacts that impair reproductive success are potentially more indicative of potential population level effects than those which may be construed as individual-level impacts (morphological abnormalities not associated with reproductive ability).

RESPONSE 3.2-RS-6

EPA concurs that impacts that impair reproductive capabilities are more indicative of potential population-level effects. Morphological abnormalities to reproductive organs as well as other tissues were evaluated and considered when characterizing risks to amphibians as described in Appendix E, Section 4.3.3.

Timothy Thompson:

No further comments here.

3.2(e) Were the statistical techniques used clearly described, appropriate, and properly applied for the objectives of the analysis?

Valery Forbes:

Generally yes. The exception here is with EPA's leopard frog study in which the control (composed of purchased frogs) was not a true statistical control.

RESPONSE 3.2-VF-9

Please refer to the response to General Issue 12.A.

Thomas W. La Point:

Yes. However, the use of log-linear models (chi-square) for tests of association are particularly sensitive to the individual cell sizes (Sokal & Rohlf, 1989; Quinn & Keough, 2002). An analogy can be made by thinking about an experiment in which the results are described as follows: "1/3 of the exposed frogs died, 1/3 of the exposed frogs lived, and the third frog got away!" The strength of the association between PCB soil/sediment concentration and the resulting reproductive or survivorship results depends not only on the number of comparisons (four ranges of PCB concentrations in this ERA), but on the number of occurrences in each "block" or "cell," as well.

RESPONSE 3.2-TL-4

EPA concurs that GE's analysis of leopard frog egg mass data was not particularly sensitive to detecting differences between treatments because of the small number of samples in each treatment.

James T. Oris:

I have the same concerns about transparency and ability to determine how statistical procedures were conducted as in previous sections.

RESPONSE 3.2-JO-5

Please refer to the response to General Issue 1.

The amphibian population modeling exercise was interesting and enlightening. However, there are issues that need to be addressed concerning the parameterization of the model. GE's critique of the model indicated that density-dependent effects can alter reproductive rates of amphibians. The EPA version of the model is set so that the only outcome is extinction of the population. Consideration of density-dependent effects may alleviate this instability in the model and allow for better predictions of population trajectories in reference and impacted areas.

RESPONSE 3.2-JO-6

Please refer to the response to General Issue 12.B.

Mary Ann Ottinger:

The information from the analytical techniques should be considered in the overall analyses in that there would be implications in potential underestimation of the MATCs.

RESPONSE 3.2-MO-15

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

Further, the life history of each species examined needs to be integrated into the models developed. For example, the wood frog has a relatively short time of potential exposure in the vernal pools; therefore, using data from this species in the field may result in a model that underestimates the risk to other amphibian species.

RESPONSE 3.2-MO-16

Detailed life history information was integrated into the model as described in Appendix E, Attachment E.3. A more thorough description of input parameters for the model will be included in the revised ERA. Also please refer to the Response to General Issue 6.

Finally, the specific analysis used in each study requires clarification, especially as the results are then integrated into the WOE.

RESPONSE 3.2-MO-17

Please refer to the response to General Issues 1 and 7

Brad Sample:

The statistic analyses appears to be generally suitable and appropriate, however transparency in approach is lacking as discussed in comment 1.e (above).

RESPONSE 3.2-BS-2

Please refer to the response to General Issue 1.

Ralph G. Stahl:

The statistical techniques were appropriate and properly applied for the objectives of the analysis.

Timothy Thompson:

Same general comment developed under the benthic infaunal section appropriate here – recommend a better description of statistical techniques used in the ERA.

RESPONSE 3.2-TT-4

EPA agrees and will include more detailed descriptions of statistical techniques throughout the revised ERA. Please refer to the response to General Issue 1.

3.2(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria?

Valery Forbes:

In my view applying a population modelling approach to integrate effects of PCBs (and other potential stressors, habitat features, etc.) on the individual-level endpoints measured can add considerable strength to the risk assessment. Such models can be particularly useful, for example, for comparing impacts on different life stages (e.g., how much of an impact on egg production would be equivalent to a given effect on adult mortality in terms of population level impact?). Such an approach could have been applied to the other receptor species, especially where the different assessment endpoints showed non-congruent response patterns.

As far as I can determine, given the way that the input parameters were chosen for the model used here, the addition of PCBs would have to increase the probability of extinction (unless the increased larval survival with PCB exposure could offset all of the modelled negative impacts). So although I was not surprised to see that the PCB cases increased the probability of decline I find myself asking, 'but how much of an increase in probability of decline is too much?'. I also found it intriguing (and non-intuitive) that if the modelled frog population was already declining, the additional impact of PCBs seemed to be less than if the population started from a stable state.

I recommend that the model be further explored, including consideration of various scenarios as well as a sensitivity analysis of model parameters.

RESPONSE 3.2-VF-10

A sensitivity analysis will be performed for the wood frog population model, and the results will be included in the revised ERA. Please refer also to the response to General Issue 12.B.

Thomas W. La Point:

In this reviewer's opinion, the determination of "high risk" to amphibians is supported by the toxicity evidence –particularly- and by the field evidence (lack of egg masses and/or adult females). In this instance, a precautionary approach would be dictated as well, for the frog species are probably less sensitive to PCB exposure than are the urodele salamanders.

The risk characterization would be enhanced (in my opinion) if the data in Figure 4.4-11 were to be put into an "SSD" format, as described above for benthic macroinvertebrates. The resulting statement of "90% [my estimate; not a calculation!] of amphibian effects endpoints responded to PCB concentrations of 10 mg/kg or less"

RESPONSE 3.2-TL-5

EPA agrees and will explore the use of SSD analysis in the revised ERA. Please refer also to the response to General Issue 1.

James T. Oris:

The overall characterization is supported and appropriate.

Mary Ann Ottinger:

The risk characterization is supported by wood frog data and by data from the surveys, but action to address the experimental flaws and subsequent reanalyses are needed.

RESPONSE 3.2-MO-18

EPA will carefully review the comments received from all of the Reviewers and undertake appropriate reanalyses and revisions to address uncertainty resulting from some of the studies used in the risk characterization.

Brad Sample:

The risk characterization for the field and laboratory studies appears to be supported by the available information and appropriate under the evaluation criteria.

The risk characterization for the tissue based-MATCs is not adequately described or presented. It is not clear from the text in Section E.4.5, nor from Table E.4-2 or Figures E.4-1 through E.4-3 upon which tissue concentration value the HQs were based. Were HQs calculated based on the mean or median tissue concentrations? This needs to be made clear. In addition, if HQs are calculated based on a summary statistic (or are done on a point by point basis, then averaged) this is inappropriate. HQs should be evaluated on a point by point basis and then presented as the frequency of samples that exceeded.

RESPONSE 3.2-BS-3

These HQs for amphibians were calculated separately for each tissue type and species, and were based on a comparison of observed residues to a LOAEL found in the literature. Please refer also to the response to General Issue 7.

Although a MATC was derived for sediment, it does not appear to have been used to provide a detailed spatial evaluation for the PSA. Such an analysis would be beneficial in identifying potential hotspots and for describing the spatial distribution of exposure and risk.

RESPONSE 3.2-BS-4

As discussed in the response to General Issue 1, expanded graphical displays of data and results will be included in the revised ERA.

Ralph G. Stahl:

The characterization of risk is not fully supported by the available information. The results of the leopard frog toxicity studies are suspect given the problems noted earlier. Field studies by EPA and GE suggest populations of leopard frogs are present in a number of pools within the PSA; however, whether those populations are numerically sub-optimal as a result of exposure to tPCBs and TEQ is not clear. In some instances it appears that habitat may be limiting factor for the presence of leopard frogs.

RESPONSE 3.2-RS-7

EPA believes the results of the leopard frog studies are sufficient to demonstrate that leopard frog populations in the PSA are under stress that is related to the presence of PCBs in soil and sediment. There is no credible evidence to suggest that these effects are related to habitat differences. For additional information, please refer to the response to General Issue 12.A.

The results of the FEL wood frog toxicity studies suggest potential impacts to metamorphs, but the case for these impacts to be ecologically significant is not well made. Further, it is not clear that these impacts are effective at the population level either.

RESPONSE 3.2-RS-8

Population modeling for wood frogs was conducted to link effects observed in FEL's wood frog site-specific toxicity study with potential impacts to local wood frog populations. EPA will include an expanded description of the process used and the results in the revised ERA. Please refer also to the response to General Issue 12.B.

The wood frog field work conducted by EPA indicates low levels of DELTs (deformities, erosion, lesions, tumors) in wood frogs within the PSA and reference area. This supports the FEL wood frog toxicity results where morphological impacts to metamorphs were observed.

However, there is no field evidence to support a conclusion that there is a significant impact to the wood frog population.

RESPONSE 3.2-RS-9

Please refer to the response to General Issue 12.B.

Timothy Thompson:

Within the context stated above, the general statements of risk from the field/lab studies appropriately support the risk statements made.

This is the first ROC where exposure modeling was incorporated, and I support the use of the models in this ERA. However, for this assessment endpoint the modeling leads to the extinction of amphibians in the PSA – which appears to stretch credibility in that the two ranid species in the PSA are "present and accounted for". One recommendation is that there be a better discussion of the parameterization of the models, and conducting a sensitivity analysis. For example, what set of parameter conditions would need to exist to not lead to extinction? Given the data collected, there should a discussion of the disconnect between model predictions and field observations.

RESPONSE 3.2-TT-5

EPA agrees that these additional aspects of the wood frog population modeling should be conducted, and will include the results in the revised ERA. For additional detail, please refer to the response to General Issue 12.B.

3.2(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made.

Valery Forbes:

The best way to address the uncertainties indicated in the field studies (due to small sample size and lack of information on actual exposure) would be to perform additional studies.

RESPONSE 3.2-VF-11

EPA believes that there are sufficient data to characterize risk to amphibians based on the site-specific toxicity studies, field surveys, and effects studies in the literature. A more thorough description of the uncertainties in the studies used to characterize risk to amphibians will be included in the revised ERA.

Thomas W. La Point:

The uncertainties were addressed in the ERA. What was not as well addressed included data on urodeles. The uncertainty here is high – but unknown. Improvements would call for data specifically on salamanders or closely related species. This would probably best be accomplished under laboratory or "mesocosm-type" controlled exposures.

RESPONSE 3.2-TL-6

EPA agrees that toxicity testing using salamanders would provide additional information for a group of species believed to be sensitive to PCBs. EPA believes, however, that there are sufficient data to characterize risk to amphibians based on the site-specific toxicity studies, field surveys, and effects studies in the literature.

James T. Oris:

Uncertainties were addressed adequately.

Mary Ann Ottinger:

Yes, there were significant uncertainties in some additional targeted endpoints, such as thyroid hormones in the Northern Leopard Frog study might correlate with the observed delay in larval development as well as low incidence of metamorphosis. Measurement of thyroid hormones would be an important measure as well as histological analysis of the delayed individuals to determine if the thyroid was abnormal.

RESPONSE 3.2-MO-19

As noted in Response 3.2-TL-6 above, additional investigation would likely add to the amphibian toxicological data base and may remove some of the uncertainties that arise from incomplete knowledge. EPA believes, however, that there are sufficient data currently to characterize risk to amphibians based on the site-specific toxicity studies, field surveys, and effects studies in the literature.

Brad Sample:

The significant uncertainties appear to have been identified and evaluated.

Ralph G. Stahl:

The uncertainties in the analysis of the amphibian assessment endpoint were identified clearly and adequately.

Timothy Thompson:

As noted previously, a global issue is that for tissue analyses the intra-lab variability discussed in Appendix C.11-1 was 28.6%. The potential effect on the interpretation of risk should be included in the final ERA.

RESPONSE 3.2-TT-6

EPA will explore methods of incorporating intra-lab variability into the risk characterizations. Please refer also to the response to General Issue 1.

3.2(h) Was the weight of evidence analysis appropriate under the evaluation criteria? If not, how could it be improved?

Valery Forbes:

Sections 4.7.1.1 – 4.7.1.3 were excellent – a clear and transparent description of the thought process going into the weighting criteria.

Apparently GE's wood frog study measured 11 endpoints but only found effects on 2 (malformations and sex ratio). However the ERA only focused on the 2 that showed effects, despite that other of the endpoints are relevant for assessing survival and reproduction. These other endpoints should be incorporated into the WOE.

RESPONSE 3.2-VF-12

All of the measurement endpoints in the GE wood frog study were evaluated in the risk assessment, although focus was placed on sensitive endpoints that could have individual and population-level effects. As with most species, certain organs, processes, or life stages are more sensitive than others to a certain COC. This relates to the mechanism of action and the opportunity for an effect to occur. The revised ERA will contain a description of current knowledge on the specific effects of PCBs on amphibians, and this will be tied into the discussion of sensitive and insensitive endpoints. Please refer also to the response to General Issue 8.

Thomas W. La Point:

For the ranid species, the WOE was appropriate. For amphibians in the floodplain, probably not. However, this could only be made more robust by collecting more information.

RESPONSE 3.2-TL-7

Please refer to Responses 3.2.TL-6 and 3.2-MO-19.

James T. Oris:

See comments in section B for WOE analysis.

Mary Ann Ottinger:

The WOE appears to be appropriate and studies were thorough. Some difficulties arose from the following: there were relatively few animals found in the initial survey [I. 4-21], whereas in the EPA survey, there were differences in species richness in number of wood frogs and in species of salamanders (I.4-64; Figure E.2-1). Reference site and purchased controls (none available from the reference sites) potentially confound the interpretation. Leopard frog data were difficult to interpret because of low fertilization success in the field collected females. Elevated sperm abnormalities in males [I. 4-36] may be an important COC indicator for these species.

RESPONSE 3.2-MO-20

Please refer to the response to General Issue 12.A. EPA concurs that the elevated occurrence of sperm abnormalities may be an important indicator of reproductive impacts from a COC.

Brad Sample:

The weight-of-evidence (WoE) analyses is generally appropriate but is not adequately transparent.

The WoE analyses use best professional judgment to apply final weightings for the various lines of evidence and their attributes. It is unclear how final weightings were derived. Some discussion as to how these weights were derived needs to be provided – were all attributes judged equally? If not need to know why and this needs to be explained – this will allow reviewers to judge the process.

RESPONSE 3.2-BS-5

EPA agrees that the transparency of the WOE weighting process could be improved, and will provide additional detail and background in the revised ERA. Please refer also to the response to General Issue 8.

Ralph G. Stahl:

The WOE analysis was not fully objective, but was reasonable. As stated previously, there is a high degree of professional judgment in the WOE analysis which tends to increase the subjectivity. And, it is not altogether consistent either as applied across all 8 of the assessment endpoints.

RESPONSE 3.2-RS-10

Although the Massachusetts WOE approach is effective in reducing the subjectivity of the WOE process, a certain amount of subjectivity and professional judgment remains. As discussed in the response to General Issue 8, EPA will include additional detail on the WOE process in the revised ERA.

Timothy Thompson:

The same general comments from the benthos comments concerning the weighting of low, moderate and high risk apply to the amphibians as well. Otherwise, the final assessment of risk to amphibians appears to be appropriate – however with a high degree of uncertainty.

RESPONSE 3.2-TT-7

For the revised ERA, EPA will re-evaluate the weighting assigned to all lines of evidence. Please refer also to the response to General Issue 8.

3.2(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?

Valery Forbes:

Yes, the landscape analysis in combination with sediment PCB concentrations seems to be a good way to do this. It is unfortunate however that there were no sediment samples available from the downstream vernal pool habitats. Taking such samples would be one way to reduce uncertainty.

RESPONSE 3.2-VF-13

EPA agrees that taking additional samples downstream would reduce uncertainty, and this may be necessary in making remedial decisions. Downstream vernal pools were not targeted specifically for sampling because the concentrations of tPCBs in downstream sediment and floodplain soil were substantially lower than those found in the PSA, and at the time of the sampling, the outcome of the risk assessment was not known. In addition, in areas downstream of the PSA in certain reaches of river in MA and the majority of CT, the frequency of vernal pools in the floodplain decreases because the floodplain is narrow and the gradient of flow is greater.

Thomas W. La Point:

Yes. As the basis for most of the risk characterization is based on toxicity studies and malformations, the use of predictive approaches downstream are appropriate. The population modeling approaches (meta-population matrixes) are sensitive, when the matrix values closely approximate the seasonal means and variances in survivorship and reproduction. Again, basing the risk characterization on ranid frog species would appear to call for a conservative basis, as urodeles are strongly suspected of being more sensitive to PCB exposure.

RESPONSE 3.2-TL-8

EPA concurs that certain amphibian species that were not studied in detail as part of the ERA may be more susceptible to the effects of PCBs because of their life history, as discussed in Appendix E, Section 4.8.2.

James T. Oris:

There is high uncertainty characterizing the downstream reaches. Apparently no samplings were conducted downstream of Woods Pond, there is little discussion of habitat suitability, and the risk estimates were based on measured/modeled PCB data.

RESPONSE 3.2-JO-7

The downstream reaches were characterized at a level necessary to develop risk estimates based on existing data. Sediment samples were collected downstream, and an ecological characterization was conducted for the river downstream of the Woods Pond Dam. This characterization (see ERA Appendix A.2) includes a review of habitat types, species expected to occur in each habitat, and expected seasonal use. In each section of the ERA, risks to species expected to be similar to representative species were described using habitat suitability and sediment sampling data. These points will be clarified in the revised ERA.

Mary Ann Ottinger:

The risk estimates are based on the GE survey in 2003 and on the Leopard frog egg mass numbers in the PSA, which did not relate to tPCBs. Based on the EPA species richness data and on the effects data, the impact of PCBs decreases at lower soil concentrations and are of concern in areas retaining high soil PCB levels.

RESPONSE 3.2-MO-21

Risk estimates for downstream areas were not based on the field studies, either GE's 2003 wood frog study or GE's 2003 leopard frog egg mass study, they were based on concentration-response analyses for certain toxicity endpoints in the wood frog toxicity tests. As described in Appendix E, Section 4.3, risk estimates were based on comparisons of the magnitude of effects for the various endpoints considered to be biologically relevant. Effects concentrations were calculated for the endpoints based on sediment and tissue tPCB concentrations and the frequency of effects. These data were used to generate MATCs, which were then used to estimate risks downstream by comparing the MATCs to measured PCB concentrations.

Brad Sample:

The same issues identified for the development and application of MATCs for benthic invertebrates (1.f, above) apply to the MATCs for amphibians. Use of the same decision process for derivation of the MATC and point by point application as for benthic invertebrates is recommended.

RESPONSE 3.2-BS-6

EPA agrees that the same procedures for evaluating the utility of data sets for MATC derivation can be applied to both Assessment Endpoints. For additional information, please refer to the response to General Issue 6.

It appears that data for PCB concentrations in downstream vernal pools are lacking. As these are the habitats in which amphibians are most likely to be exposed and affected, the lack of vernal pool data from downstream reduces strength of downstream analyses. This lack of data and its effects should be noted.

RESPONSE 3.2-BS-7

The lack of data for vernal pools downstream of the PSA, and the reasons EPA believes these data are not necessary for evaluating potential risks to amphibians, will be discussed in the revised ERA. Please refer also to Response 3.2-JO-7.

Ralph G. Stahl:

The risk estimates were derived objectively and appropriately for reaches of the river where site-specific studies were not conducted. However, as noted previously, the MATC for sediments in the PSA was not based on the most synoptic data and is therefore overly conservative. As the basis for estimating risks to other reaches of the river where studies were not conducted, then it may also be providing an overly conservative estimate.

RESPONSE 3.2-RS-11

The amphibian MATC was derived using the most synoptic data. EPA does not agree that this MATC is overly conservative and believes that it can be used appropriately to extrapolate risks downstream of the PSA. Please refer to the response to General Issue 6.

Timothy Thompson:

I concur with Panel colleagues who pointed out that the absence of samples from the vernal pools downstream of the PSA limits the ability to make statements about relative risks to amphibians.

RESPONSE 3.2-TT-8

EPA does not agree that data from vernal pools downstream of the PSA are necessary to estimate risks in downstream reaches. Please refer to Response 3.2-JO-7.

Also, the MATC criteria developed for benthos are equally applicable to amphibians.

RESPONSE 3.2-TT-9

EPA agrees that similar criteria for data sets used for MATC derivation apply to both benthos and amphibian endpoints. Please refer also to the response to General Issue 6.

3.2(j) In the Panel members' opinions, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?

Valery Forbes:

The ERA concluded that risk to amphibians is high and that confidence in this conclusion is high. Although I agree that the probability of some effects occurring in amphibians is high, it is not as clear to me that the magnitude of these effects is high.

RESPONSE 3.2-VF-14

Wood frog population modeling was conducted to translate the individual-level effects observed in site-specific toxicity testing to population-level effects. Modeling results indicated that the potential magnitude of effects to amphibians resulting from tPCB contamination are high. The discussion of the modeling results and the potential magnitude of effects will be expanded in the revised ERA. Please refer also to the responses to General Issues 12.A and 12.B.

Thomas W. La Point:

Yes, in part. As a matter of professional judgment, the potential for sediment toxicity to larval frogs should be "moderate to high," given the high weighting, the demonstrated evidence of harm (albeit from laboratory toxicity studies) and the fact that the frogs are surrogates for salamanders, strongly suspected of being more sensitive to PCB exposure.

RESPONSE 3.2-TL-9

EPA will review the weighting criteria for all of the Assessment Endpoints. For additional general discussion of the WOE process, please refer to the response to General Issue 8.

James T. Oris:

Within the PSA, yes. However, further discussion of downstream reaches is needed.

RESPONSE 3.2-JO-8

Please refer to Response 3.2-JO-7. Additional clarification will be added to the ERA to more fully explain how risk estimates downstream were developed for amphibians.

Mary Ann Ottinger:

The data support the conclusions in spite of some shortcomings in the data set.

Brad Sample:

Based on the information provided in the ERA, the risk conclusions are probably accurate and appropriate.

Ralph G. Stahl:

The conclusion of high risk to leopard frog populations is not fully supported by the information provided in the ERA. There are significant questions regarding the conduct of the FEL toxicity studies on the leopard frog, and the timing and conduct of observations in the field with respect to leopard frog egg masses.

RESPONSE 3.2-RS-12

A more complete description of the uncertainties associated with FEL's leopard frog study and the relationship to GE's leopard frog egg mass study will be presented in the revised ERA. EPA does not believe that there is question regarding the timing and conduct of observations of leopard frog egg masses in the field made by either EPA or GE. Please refer also to the response to General Issue 12.A.

The conclusion of high risk to wood frog populations is partially supported by the information provided in the ERA. The FEL toxicity studies on the wood frog indicate potential morphological effects on young life stages from exposure to tPCBs and TEQ, yet these effects are not evident in any widespread nature from the field observations in the PSA (low levels of DELTs). In my opinion, the more defensible conclusion of risk to wood frog populations is that it is moderate, given the lack of concordance between the field and laboratory studies.

RESPONSE 3.2-RS-13

As described in Section E.4.3.1.1, there was concordance between sensitive endpoints in the field and laboratory studies. Field observations of DELTs include Phase III metamorphs where there were internal and external malformations that were significantly correlated with both tPCB in sediment and in tissue. Rates of gross external DELTs in adult frogs captured entering vernal pools to breed were, however, low everywhere and not significantly different between contaminated and non-contaminated pools. There are several reasons for this finding, one being that most frogs with gross external DELTs do not survive to breed in vernal pools.

Timothy Thompson:

In general, within a high degree of uncertainty, yes. However, I do not believe these data contain sufficient strength to formulate a supportable MATC around which a risk management decision can be constructed.

RESPONSE 3.2-TT-10

EPA believes that there are sufficient data, including site-specific toxicity tests, to develop a supportable MATC for amphibians. Appendix E, Section 4.3, includes a detailed analysis of the process used to develop effects concentrations for different endpoints using both sediment and tissue tPCB concentrations. There were statistically significant correlations with sediment and tissue tPCBs for larval malformations, metamorph malformations, gonadal malformation, and skewed metamorph sex ratios. There was also agreement between the site-specific toxicity tests, the wood frog laboratory study (Phase I), and the field component (Phase III). The findings of these studies indicated that PCBs were readily taken up by frog larvae, and that they were causing malformations and skewed sex ratios at statistically significant levels. There was additional confirmation from the findings in the crossover and spiking studies. EPA believes that this provides a very robust set of data on which to base the development of an MATC and risk characterization. Please see also Response 2-RS-5 and the response to General Issue 6.

3.3 Fish

> ***3.3.(a) Were the EPA studies and analyses performed (e.g., field studies, site-specific toxicity studies, comparison of exposure and effects) appropriate under the evaluation criteria, and based on accepted scientific practices?***

Valery Forbes:

Neither the EPA nor the GE field studies were optimally designed to test concentration response relationships. However both studies seemed appropriate for assessing the condition of fish populations in the PSA and therefore contribute important information.

Thomas W. La Point:

Yes. Electroshocking and nest surveys are readily accepted as methods. Further, the studies took pains to ensure that "sampling-effort" was standardized among the different field collections. The literature data, summarized in Fig. 5.3-1, is an excellent example of intensive literature-based analysis. With the results explained in Figs. F.4-7 and F.4-8, there is a clear indication of the TEQ threshold of 45 – 50 ng/kg and the 30 – 45 mg/kg threshold for tPCB.

James T. Oris:

In general, studies and analyses were done as described and meet accepted scientific practice. However, it is my opinion that the assessment endpoints were too restrictive for fish in this assessment.

RESPONSE 3.3-JO-1:

Please refer to the responses to General Issues 2 and 13.E.

Presence of disease and deformities should be used as indicators of population health, but were not apparently used in this way.

RESPONSE 3.3-JO-2:

Please refer to the response to General Issue 13.E. There were differences of opinion among Reviewers regarding the relevance of disease and deformities (and other individual responses such as endocrine and histological effects) as measures of population health. EPA agrees that individual-level responses should be discussed in the risk characterization, but that these should be placed in the context of measures of population effects.

Studies should also have been done to assess population structure and growth modeling (as with amphibians), but were not.

RESPONSE 3.3-JO-3:

Please refer to the response to General Issue 13.F. EPA agrees that a discussion of the population structure would contribute to a better understanding of the effects on the local population, and will provide a more detailed discussion in the revised ERA.

Insufficient data were collected to determine these higher-level effects.

There were some deficiencies in the dose response relationships in the egg injection studies that need to be addressed. Examination of the data in attachment F.7 indicates that there were problems with the Largemouth bass studies. There did seem to be a clear dose response, however, using the rainbow trout model.

RESPONSE 3.3-JO-4:

EPA agrees that there is more uncertainty associated with the data from the largemouth bass studies than the rainbow trout study. Mortality of largemouth bass (LMB) was high at 15-days post swim-up, which was believed to be related to the failure of the bass to survive the transition from endogenous to exogenous feeding. For this reason, the 15-day post swim-up data did not meet quality control criteria, and were not subject to further statistical analysis. However, the data for survival at swim-up were adequate for statistical analysis. Although the variability in dose-response curves was greater for largemouth bass than rainbow trout, the data were adequate for dose-response evaluation. Furthermore, because the manifestation of PCB-related responses is expected to be greatest during the period of yolk-absorption, the responses observed at swim-up in LMB may underestimate the toxicity that would be observed had the study been successfully carried through yolk-sac absorption. EPA will incorporate these considerations in the uncertainty assessment in the revised ERA.

With regard to dose-response relationships, EPA acknowledges that not all trials for all species or reaches exhibited a linear or sigmoidal dose-response relationship, but such responses would not be expected to occur consistently in a study of this type. Certain responses to PCBs are not well predicted with a simple model. The responsiveness of any endpoint is related to the level of biological organization (molecular, biochemical, histological, etc.) and the mode of action of the contaminant, and varying sensitivity of individuals.

Primary responses to PCBs directly coupled to the aryl hydrocarbon receptor (AhR) typically demonstrate a more predictable dose-response. Secondary effects from exposure to PCBs, such as reductions in estradiol, have more variables influencing the dose-response relationship and thus will be less predictable from simple models. At each increase in the biological level of organization, from molecular to biochemical to tissue, and ultimately to population-level responses, there are an increasing number of variables that contribute to the response. Therefore, it is not expected that a simple model would predict the higher-level responses with the same level of precision or accuracy as lower-level responses.

In addition, the natural variability of the test organisms is sufficiently high that interrupted dose-response relationships in some trials would be expected, even if the exposure term were measured without error. The relatively small PCB and TEQ concentration differences across exposed treatments, combined with the potentially non-linear form of the dose-response relationship, increase the probability of interrupted dose-response. Rather than focusing on individual trials, EPA adopted a broader interpretation that examined the comprehensive patterns of responses, an approach used by the Principal Investigator (Dr. Donald Tillitt, USGS), based on his experience with experiments of this type. EPA acknowledges that any departures from continuous dose-response relationships should be discussed in the revised ERA. Please refer also to the response to General Issue 13.A.

Mary Ann Ottinger:

As with other wildlife species, some species of fish move within the PSA and may receive varying exposure to contaminants. Therefore, the identification of classes of fish and the selection of representative species within each is an excellent approach. The Phase I Largemouth bass study showed a number of endocrine and histological end points relevant for reproductive impact, including reduced steroid hormones and gonadal abnormalities. These are reliable and appropriate measurement endpoints for reproduction; however these data are not provided in the tables or figures.

RESPONSE 3.3-MO-1:

Please refer to the response to General Issue 13.E.

EPA agrees that endocrine and adult histological endpoints should be presented in more detail. In the revised ERA, EPA will consider both the magnitude and consistency of the observed responses, and include a discussion of the linkage to population health. Histological responses, such as the gonadal abnormalities observed in the Phase I adult largemouth bass, are complex higher-level responses that are less likely to be precisely predicted by a simple dose model.

The field studies (EPA and GE) assessed species abundance, density, and largemouth bass reproduction and population, respectively.

Brad Sample:

The EPA studies appear to be suitable and appropriate.

Ralph G. Stahl:

The EPA field studies were appropriate under the evaluation criteria and were based on accepted scientific practices.

The site-specific toxicity studies and comparison of exposure and effects were appropriate under the evaluation criteria and based on accepted scientific practices. However, there were problems noted in the egg injection studies where there was a high percentage of the eggs that became impaired due to the injection of the solvent control substance. This suggests that there were egg handling problems associated with this study, which increases the uncertainty associated with any conclusions drawn from this portion of the toxicological effort.

RESPONSE 3.3-RS-1:

EPA contacted the Principal Investigators regarding this issue, and they indicated that the data do not suggest that there is a concern due to egg-handling effects. Nevertheless, EPA will include a discussion of comparative responses between the negative control and the solvent control in the revised ERA. The method of handling controls in the Phase II study protects against spurious outcomes resulting from solvent injection. Specifically, if results in the two controls were found to be significantly different, the solvent control was used to make corrections, whereas the controls were averaged when no statistically significant difference was observed. This procedure is consistent with the guidance found in ASTM (2004) (Section 9.2.4.3):

> If the test contains both a dilution-water control and a solvent control, the survival and growth of the organisms in the two controls should be compared (see X10.6). If a statistically significant difference in either survival or growth is detected between the two controls, only the solvent control may be used for meeting the requirements of 13.1.8 and as the basis for calculation of results. If no statistically significant difference is detected, the data from both controls should be used for meeting the requirements of 13.1.8 and as the basis for calculation of results.

Timothy Thompson:

The EPA studies are solid investigations that are consistent with, or are on the cutting edge, of accepted scientific practice. This includes the fish egg exposure/toxicity studies conducted by Tillitt et al (2003), the fish biomass estimates (Woodlot 2002), and the numerous studies on fish tissue PCB levels that have been performed for the RCRA Facilities Investigation.

Specific issues that I have with these studies relate to (1) presentation of lipid-normalized data, (2) quality assurance/quality control issues relating to PCB and lipid measurements, (3) interpretation of exposure and effects, and (4) data omission/inclusion for the MATC calculation.

The presentation within the ERA makes the implicit assumption that PCBs are directly tied to lipid content without making that demonstration. While the literature argument for PCB "normalization" is made within the text, that confirmatory analysis is not made and should be considered. Distinct differences have been noted for different species and collection season between PCBs and lipid concentration in fish. In analysis of time trends for the Lower Fox River and Green Bay, a given species and sample type differed between the reaches in one or more ways: 1) average PCB concentration differed, 2) time trend in PCB concentration differed, or 3) the relationship of PCB concentration to lipid content differed (TMWL and RETEC 2001). PCB associations with lipids were found to vary according to species, season collected, and area within the River or Bay PCB concentrations varied with percent lipid in different ways within two River reaches for two species of fish. Lumping across species does not take into account differences that have or have not occurred due to the well-known phenomenon of material transfer of lipid and contaminants between females and their eggs. For the Housatonic, the normalization may be appropriate – however it has not been demonstrated and needs to be done.

RESPONSE 3.3-TT-1:

Please refer to the response to General Issue 13.C.

Site-specific information on the relationship between lipid content and bioaccumulation in PSA fish indicates that lipid content is one of several important controlling factors for PCB bioaccumulation. This is supported by the bioaccumulation model results (WESTON 2004) and field data (WESTON 2003). An independent evaluation by GE of adult largemouth bass PCB concentrations and lipid contents (BBL and QEA, 2003) confirmed the findings of the EPA fish tissue collections. This study provides some evidence that the relationship between PCB and lipid content is stable across sampling years.

EPA will summarize this information in the revised ERA to provide a more detailed rationale for lipid normalization.

The occurrence of low lipid values (< 0.3%) in fish tissue samples is discussed in Appendix D to the RFI, but are not accounted for in the ERA. According to that appendix, this is especially true for fish fillet tissue samples, and that for some species these account for a large percentage of the existing data. This appears to be especially true for largemouth bass, yellow perch, and brown bullhead within Reaches 5 and 6. These data are used in the ERA, and a formula for deriving the whole body concentration from fillet data are presented in Section F.2.2, but lipid data are not discussed. As correctly noted in the RFI,

> "When dividing similar fillet PCB concentrations by their corresponding lipid contents, the lipid values that are low by a factor of 10 will result in normalized PCB concentrations that are skewed high relative to the remainder of the data"

In light of the importance lipid normalization is given in the ERA, this uncertainty should be addressed and discussed.

RESPONSE 3.3-TT-2:

EPA agrees that discussion of the uncertainty associated with low lipid concentrations should be included in the ERA. However, EPA does not believe that the effect of this uncertainty on the characterization of risk for fish is large. The majority of samples with low lipid content identified by GE were fillet samples. The whole body fish tissue PCB concentrations depend mainly upon the concentrations in the offal portion of the fish. This is due to the greater mass of offal samples relative to fillet (approximately four times greater, on average). Furthermore, an independent study of largemouth bass by General Electric in 2002 (15 samples each in Reach 5B and Reach 6) indicated whole-body lipid contents that were in agreement with the EPA data. A quantitative comparison will be incorporated in the revised ERA.

The results of the fish studies are not presented in an objective fashion. For example, the text on Page F-31 discusses a range of effects that were observed in largemouth bass offspring in the PSA ("which were statistically different from those observed in the offspring from the reference location). These differences included reduced survival from hatch to swim up, development delays (increased days to swim-up), growth, and deformities. What is not covered is a discussion of how those related to tPCB concentrations reported in the whole adult largemouth bass in the ovaries.

RESPONSE 3.3-TT-3:

Please refer to the response to General Issue 13.A.

While there is a thorough discussion of the fish toxicity studies and endpoints provided in the USGS reports (attachments to the ERA), EPA will provide a more detailed evaluation of these endpoints directly in the revised ERA. However, it should be noted that to demonstrate adverse effects, it is not necessary to demonstrate dose-response relationships for all endpoints. The number and type of endpoints exhibiting effects should be compared against the number and type of endpoints that do not show effects, as part of a lines-of-evidence approach.

There was disagreement among individual Reviewers regarding the ecological significance of some individual-level biological responses (particularly in terms of population-level impacts). The revised ERA will include more discussion of biomarkers, such as endocrine and histological endpoints, and will also point out endpoints that showed weak or non-existent concentration-response relationships. This will make the effects assessment more balanced.

That tPCB effects on fish are observed is not in dispute; rather it is the "selective" presentation of information that is called into questions. The table below presents three of those parameters versus tPCB concentrations from the Tillitt Phase 1 study. In the case of survival, there is clearly NO dose/response relationship; this is not discussed in the text. Rather, the ERA presents the negative effects observed without placing into context the observations. GE appropriately notes this within their comments, and I concur; as presented dose/response is not demonstrated. This does not meet standards for presentation of scientific information, and should be corrected. This also continues over to the Phase 2 study (e.g., craniofacial abnormalities in injected largemouth bass).

RESPONSE 3.3-TT-4:

EPA believes that the data demonstrate a dose-response relationship for Phase I survival. The survival to swim-up stage was significantly reduced in all three exposed reaches relative to the reference station. The range of exposures within PSA reaches (i.e., less than two-fold difference in ovary tPCB concentrations) was sufficiently small that a monotonic dose-response relationship was not expected. As discussed previously, the 15-day post swim-up data did not meet quality control criteria, and were excluded from further statistical analysis. The lack of dose-response observed for that endpoint, therefore, should not be interpreted as an indication of lack of biological response, but rather an inability of bass to switch to exogenous feeding in the laboratory that negates the ability to analyze these data.

EPA will include an expanded discussion of the analysis of Phase I and II data in the revised ERA.

tPCB Concentrations and Hatchout Parameters for LMB

Tillitt Sample ID	Reach	Concentration tPCBs in Ovaries (Table 25 Mean ng/g ww)	Concentration tPCBs in Whole Bass (Table 24 Mean ng/g ww)	Survival to Swim-Up (Table 11 Mean)	Days from Hatch to Swim Up (Table 12 Mean)	Survival 15 days Post Swim-up (Table 15 Mean)
Threemile Pond	Reference	634	106	68	7	28
Rising Pond	8	193,000	43,000	45	8	2
Woods Pond	6	251,000	108,000	57	9	26
Deep Reach	5C	315,000	149,000	28	8	13

3.3(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria, based on accepted scientific practices, and incorporated, appropriately in the ERA?

Valery Forbes:

See response to 3.3.a.

Thomas W. La Point:

Yes.

James T. Oris:

Studies conducted by GE to assess the presence of spawning activity and natural reproduction of fish in the PSA were appropriate given the narrow scope of the study design, used accepted scientific methods, and were incorporated appropriately into the ERA. However, no attempt was made to correlate levels of contaminants and these measures, and (in the ERA) there is no analysis of the amount of potential habitat occupied.

RESPONSE 3.3-JO-5:

EPA agrees that the GE field studies are limited by the lack of evaluation of contaminant concentrations in the study. EPA also agrees that determining the amount of potential habitat occupied was not a specific objective of the GE study.

In addition, there is little discussion of other sources of recruitment (e.g., non-contaminated tributaries or contaminant "cold"-spots in the PSA).

RESPONSE 3.3-JO-6:

The GE largemouth bass study evaluated both main channel habitat and non-contaminated tributaries, and concluded that the latter were incapable of supporting significant reproduction due to unsuitability of habitat. Although EPA has concerns with some technical aspects of the R2 study, EPA supports the general conclusion that tributaries are not a dominant source of recruitment to the bass population observed in the main channel.

The chemical characterization of the sediment in the PSA does not indicate large areas that have consistently low PCB concentrations (i.e., “cold” spots) except perhaps in some of the backwaters removed from the river channel. There is significant variation in PCB concentrations in individual fish, which may be partly related to small-scale variations in sediment concentrations. However, the data do not exist to support a conclusion that reproduction occurs only in “clean” areas of the PSA.

Mary Ann Ottinger:

The GE study in bass did not relate contaminant concentration to response. It is interesting that the adults were generally older. This may only be a sign of year to year differences in the hatching and survival of young of the year (YOY) or it may reflect survivorship of a subpopulation that are more resistant to the COCs. Because PCBs have been in the environment for a long time and bass are long-lived fish, it is likely that the older adults are survivors of the environmental challenge.

RESPONSE 3.3-MO-2:

Please refer to the response to General Issue 13.F.

Brad Sample:

GE study appears to have been well designed and conducted. The results of this study however do not appear to have been correlated to the PCB concentrations, so not dose-response association can be made. Data indicate that the population is skewed toward older fish – this could indicate adverse effects on recruitment. Overall, the study was correctly used by EPA.

RESPONSE 3.3-BS-1:

The Reviewer is correct that the GE study did not include an evaluation of PCB concentrations. In addition, please refer to the response to General Issue 13.F.

Ralph G. Stahl:

The GE field studies for largemouth bass were appropriate under the evaluation criteria, based on accepted scientific practices, and were incorporated appropriately in the ERA. The study design was specific to understanding potential habitat related influences on the population of largemouth bass residing in the PSA.

RESPONSE 3.3-RS-2:

EPA does not agree that the design of the GE largemouth bass study was specific to understanding the effects of “habitat related influences on the population of largemouth bass residing in the PSA.” The objectives of the study were to determine whether the largemouth bass population is self-sustaining, whether the population depends on tributary recruitment, and whether population parameters are similar to other systems. Although data were collected that identified areas of habitat that were either suitable or unsuitable for largemouth bass reproduction, the study was not specifically designed to quantitatively link habitat parameters to the numbers of bass present or reproductive output.

Timothy Thompson:

The GE analysis is a very comprehensive picture of bass ecology and reproduction within the PSA. However, there study focused solely on demonstrating that bass spawning areas do exist within the PSA, and generally demonstrating that spawning occurs. That is very different from reproductive impairment, which is what EPA studies are intended to demonstrate. I concur with Dr. Oris the GE made no attempt to correlate the tPCBs in or near the nests. It's possible that since these areas were principally at the mouths of tributaries to the Housatonic, tPCBs would not accumulate there. A more convincing point would be to measure tPCBs within the sediments at the spawning areas. Having said that, egg "exposure" even in these areas will come from maternal transfer – not uptake from local sediment sources.

RESPONSE 3.3-TT-5:

EPA agrees that the GE study did not evaluate all aspects of reproductive impairment, and also agrees that the study demonstrates that spawning occurs in the Housatonic River. However, the limited measurement endpoints (i.e., no direct measure of reproduction or recruitment) and lack of dose-response analysis do not enable a definitive conclusion to be drawn concerning whether recruitment is adversely impacted by PCBs, or whether the resilience of the population to additional stressors would be affected. Only very large disruptions to spawning and reproduction behavior could be detected given the GE study design.

3.3(c) Were the estimates of exposure appropriate under the evaluation criteria, and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?

Valery Forbes:

Mapping of exposure of fish populations in space would be a very useful addition; i.e., where were fish tissue data collected?

RESPONSE 3.3-VF-1:

Please refer to the response to General Issue 1.B; maps will be provided showing locations of fish collections.

However, it is recognized that f or some COCs fish tissue would not be a good measure of exposure.

Thomas W. La Point:

Yes.

James T. Oris:

Estimates of exposure and the COC refinement seemed appropriate, but EPA needs to recognize that COC's other that PCBs can cause many of the same effects in fish. A better accounting of the other COC's needs to be present. The Weston "in prep" document should be included in the ERA documentation.

RESPONSE 3.3-JO-7:

EPA agrees that other contaminants can elicit adverse responses in fish, including some of the responses observed in the fish toxicity studies. The Reviewers have specifically identified mercury and PAHs as potential causative agents in the EPA studies; however, it is highly unlikely that these contaminants exerted significant responses in the site-specific toxicity studies.

PAHs are not expected to confound the fish toxicity results for several reasons:

- The processing methods of the tissue extracts used in the Phase II study (described in the December 2003 responses to initial Reviewer questions) would reduce toxicity associated with contaminants other than hydrophobic organic contaminants (e.g., PAH concentrations reduced in processing).
- The route of exposure simulated in Phase II (maternal transfer) is not conducive to the elicitation of PAH effects.
- The offspring in both Phase I and Phase II were reared in a laboratory environment in which PAH concentrations in the water column were low. The exposure paradigms applied by USGS are inconsistent with PAH toxicosis.

Mercury was eliminated from consideration as a tissue-based COPC in the Tier I Pre-ERA (Section B.4.2), although it was identified as a sediment-based COPC in some reaches. In the receptor-specific screening stage for fish, PSA tissue concentrations were also compared to tissue concentrations measured in the least-contaminated water bodies in MA (Section 3 in Attachment F.1). Evaluation of mercury was, therefore, not necessary in the ERA. To further address concerns of the Reviewers, an evaluation of mercury toxicity to fish was conducted. The results of this evaluation are summarized below.

- Mature Fish - Wiener and Spry (1996) present a summary of total mercury concentrations in tissues of fish exhibiting symptoms of methylmercury toxicity. The lowest observed adverse effect level (LOAEL) for brook trout in a study by McKim et al. (1976) was 5 mg/kg in muscle tissue and 3 mg/kg in whole body. Effects in other studies (emaciation, enfeeblement, growth reductions, mortality increase, etc.) were often observed in the 10 to 30 mg/kg range in fish muscle. Overall, field studies indicate that fish with residues in muscle in the range of 6 to 20 mg/kg wet weight are adversely affected. The range for laboratory studies is similar, with sublethal effects or mortality associated with muscle tissue concentrations of 5 to 8 mg/kg in walleye and 10 to 20 mg/kg in salmonids. For salmonids, a concentration of 5 mg/kg in muscle can be regarded as a no-observed effect concentration, based on the work of McKim et al. (1976).

- Early Life-Stage Fish - Another endpoint of concern for mercury in fish is embryotoxicity. In wild fish, a small quantity of mercury is transferred from the female to eggs, and the sensitivity of embryos to egg concentrations of mercury is greater than that of adult fish (on a wet weight basis). Sublethal and lethal effects on fish embryos are associated with mercury residues in eggs that are much lower (perhaps 1% to 10%) than the residues associated with toxicity in adult fish. However, this consideration is offset by the fact that egg concentrations tend to be much lower than adult tissue concentrations, because only "a small quantity of mercury is transferred from the female to the eggs during oogenesis" (Wiener and Spry 1996). Niimi (1983) found mercury concentrations in eggs of five Great Lakes fish that ranged from only 0.3% to 2.3% of the maternal concentrations. Therefore, these considerations offset each other, such that the protective muscle tissue concentration is similar for adult and embryotoxicological endpoints.

In summary, the concentrations of mercury found in Housatonic River fish (i.e., less than 1 mg/kg) are below concentrations demonstrated to produce ecological effects in fish. Lethal or sublethal effects due to mercury have not been observed in the range of fish concentrations measured in the study area, in either laboratory or field studies.

EPA provided a receptor-specific screening for fish that excluded pesticides from further consideration in the ERA. PAHs were retained in the fish ERA and were evaluated via consideration of sediment PAH concentrations.

Please refer also to the response to General Issue 3 (Screening of COPCs and Determination of COCs).

Mary Ann Ottinger:

The field survey includes a number of species and residue analysis of appropriate n of animals. Gut contents were not analyzed, so the foraging area is assumed according to the species. There appears to be sufficient data to estimate the COCs in the fish. As mentioned above, the bass in the population were older individuals, suggesting lower sensitivity of these individuals and potentially different metabolic characteristics associated with the liver enzymes and oxidative metabolic processes.

Brad Sample:

The exposure evaluation appears to be suitable and appropriate. Tissue concentration data provide the most accurate measure of exposure. Use of sediment data to evaluate exposure for PAHs is suitable due to the metabolism of these compounds.

Ralph G. Stahl:

With one exception, the estimates of exposure were appropriate under the evaluation criteria, as was the refinement of analyses for COCs. The MATC for coldwater species, extrapolated from that of warmwater species, is not fully supported by the discussion in the ERA nor in subsequent responses from EPA and its contractors during the January 2004 public meeting. For example, there is little evidence to help determine whether or not the division of the warmwater MATC (49 mg/kg) by 4 is over or under protective. Thus the extrapolation of exposure and the derivation of an HQ based on a MATC divided by 4 is not objective nor consistent. It may be reasonable however.

RESPONSE 3.3-RS-3:

Please refer to the response to General Issue 13.B.

Timothy Thompson:

No. There is a discontinuity between the screening procedure applied in Appendix F, and additional measurements made under the Phase 1 Tillitt et al study that should be bridged. Section F.2.1 discusses the use of sediment and water quality values for screening out metals and pesticides as COC for fish. Given that pesticides in whole bass and ovaries were included in the Phase 2 study, those should be included in the screening assessment.

RESPONSE 3.3-TT-6:

EPA will incorporate the fish tissue pesticide information in the screening procedure in the revised ERA. Tissue pesticide data will be compared against tissue residue thresholds identified in the literature.

The statement within the ERA that PAHs in fish were not evaluated because these are metabolized is not an effective use of data or assessment of risk to fish. It is precisely those metabolic products that produce toxicity in fish species. Furthermore, it is inaccurate to state that PAHs are not measured in fish species. With sufficient exposure, carcinogenic PAHs have been measured in fish tissues. Better use of the PAH data should be made within the ERA.

RESPONSE 3.3-TT-7:

EPA agrees that toxic by-products of PAH metabolism can elicit toxic responses. The ERA did not conclude that PAH effects in fish were not worthy of consideration, but instead indicated that measurement of tissue parent PAH concentrations in fish is not a reliable means of evaluating risk. Many PAHs are capable of inducing the Mixed Function Oxygenase (MFO) system. The MFO system is responsible for the rapid metabolism and excretion of PAHs, and therefore limits bioaccumulation. Published literature (Field 1996, Varanasi et al. 1989) confirms EPA's conclusion that fish tissue residue concentrations of parent PAH compounds do not provide a useful measure of exposure to fish.

Neither parent PAHs nor metabolic byproducts were measured in PSA fish. Therefore, there is no way to make better use of these (nonexistent) data. A sediment-based PAH effects evaluation for fish was conducted and included in the ERA.

3.3(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria?

Valery Forbes:

The measurement endpoints used in the Phase I and II toxicity studies were appropriate, however linking them to impacts on fish populations is more problematic.

Some of swim bladder abnormalities seem to disappear with age. This issue needs further consideration.

RESPONSE 3.3-VF-2:

EPA will provide additional information on the timing of swim bladder abnormalities in the revised ERA. EPA notes that a delay in swim bladder inflation may still be ecologically relevant. Also, the measurements of swim bladder abnormalities at the later time period represent rates for the surviving fish only (i.e., they do not consider the mortality of fish from the previous time period). Please refer also to the response to General Issue 13.A.

Phase I spawn success data (number of spawns evaluated for abnormalities) have small sample sizes; and no clear dose-response. I recommend including only effects that show a dose-response.

RESPONSE 3.3-VF-3:

Please refer to the response to General Issue 13.A. EPA believes that the data demonstrate a dose-response relationship for Phase I survival when the entire range of exposures is considered. The survival to swim-up stage was significantly reduced in all three exposed reaches relative to the reference station. The range of exposures within PSA reaches (i.e., less than two-fold difference in ovary tPCB concentrations) was small. Therefore, given the variability in responses observed, it is not expected that a monotonic dose-response would be observed.

The sample sizes (both spawns and individual fish) in the Phase I study will be discussed in more detail in the revised ERA.

In general, care needs to be taken when basing effects estimates on the surviving portion of the population especially if survival was very low and/or variable among treatments.

RESPONSE 3.3-VF-4:

EPA concurs with this comment, which is relevant to Response 3.3-VF-2, above.

Thomas W. La Point:

Yes.

James T. Oris:

The effect metrics that were identified and used were appropriate under the evaluation criteria. However, there needs to be discussion of effects metrics that were not identified or used. For example, there appears to be lower rates of natural reproduction of fishes and higher rates of significant deformities in the PSA compared to reference areas. Effect metrics for these areas should have been identified and analyzed.

RESPONSE 3.3-JO-8:

> Please refer to the responses to General Issues 13.A and 13.E. EPA will provide more information on the reproduction surveys from the GE largemouth bass study (e.g., catch-per-unit-effort data), and incorporate information on adult deformity rates in the revised ERA.

In addition, community structure was not identified as an effect metric, but was used in the assessment.

RESPONSE 3.3-JO-9:

> Please refer to the response to General Issue 13.F. EPA agrees that population endpoints were considered in the ERA and will expand upon the presentation of population endpoints in the revised ERA, based on Reviewer comments.

In addition to identifying appropriate metrics, additional discussion or consideration of the level of the metrics used is merited. When asked why the 50% effect level was used to set toxicity benchmarks, I was told it was non-conservative and it was the "most accurate" estimate of toxicity (EPA answers to Oris written questions). However, in a toxicity frequency distribution, the 50% endpoint is the most precise endpoint estimate of toxicity, not the most accurate. The most precise endpoint provides smaller bounds of error around the endpoint estimate. The most accurate endpoint is determined by how the distribution is modeled. Thus, there seems to be some basic confusion about the choice of the endpoint. In addition, I would argue that multiple endpoints should be used (e.g., 20% and 50% as invertebrates, or the use of slope and center point of the dose response relationship) in the toxicity metric. A paper by Oris and Bailer (Environ. Toxicol. Chem.: 16: 2204-2209. 1997.) describes the rationale for this statement.

RESPONSE 3.3-JO-10:

> Please refer to the response to General Issue 1.C. EPA agrees that the calculation of endpoints for fish should consider multiple effect sizes, consistent with other receptor groups.
>
> Although the precision of point estimation for other effect sizes will be lower than for the 50% effect levels, EPA will consider the methods available for making estimates of lower effect sizes. The uncertainty in these estimates will be described in the uncertainty assessment section of the revised ERA.

Mary Ann Ottinger:

Basic reproductive end points were measured, including eggs and nests. These are not necessarily the most reliable or sensitive measures. In addition, measures of health, including morphological abnormalities should be incorporated into the assessment even though individuals survive with these deformities.

RESPONSE 3.3-MO-3:

Please refer to the response to General Issue 13.E. EPA agrees that the reproductive endpoints were not sensitive to subtle alterations in fish reproduction and recruitment; this was reflected in EPA's incorporation of the study results in the risk characterization. There were differences of opinion among Reviewers regarding the relevance of disease and deformities (and other individual responses such as endocrine and histological effects) as measures of population health. EPA agrees that individual-level responses should be discussed in the risk characterization, but that these should be placed in the context of measures of population effects.

Brad Sample:

The effects data generally appear suitable and appropriate – However, as this is out of my area of expertise, I need to defer to the other panel members on this topic.

Ralph G. Stahl:

The effects metrics used were appropriate under the evaluation criteria.

Timothy Thompson:

No. As noted above the ERA presents PCB "effects" without noting that those did not display dose/response. The absence of this discussion limits the utility of these data – and should be rectified.

RESPONSE 3.3-TT-8:

EPA believes that the data demonstrate a dose-response relationship for Phase I survival when the entire range of exposures is considered. Survival to swim-up stage was significantly reduced in all three exposed reaches relative to the reference station. The range of exposures within PSA reaches (i.e., less than two-fold difference in ovary tPCB concentrations) was small. Therefore, given the variability in responses observed, it is not expected that a monotonic dose-response would be observed. In addition, please refer to the response to General Issue 13.A.

The mixing of endpoints in the MATC calculation, and the exclusion of other data, also limit the utility of this information.

RESPONSE 3.3-TT-9:

EPA believes that combining significant pathologies in the Phase II study was appropriate; and used this approach after consultation with Dr. Tillitt, the Principal Investigator. There is not a definitive biological or toxicological foundation, or mechanistic understanding of the exact mode of action of PCBs, to allow separation of individual endpoints. All of the fish pathologies are considered to be signs of contaminant-induced stress, and that stress may be manifested differently in different trials due to natural variability. All significant pathologies were aggregated into an overall response that could be compared across trials and treatments. Because the aggregation included all pathologies, rather than only those that demonstrated dose-response, the analysis was not biased toward the identification of an overall dose-response.

EPA will review the available data and ensure that all appropriate data, including adult DELT data and histology/endocrine data, will be discussed in the revised ERA.

3.3(e) Were the statistical techniques used clearly described, appropriate, and properly applied for the objectives of the analysis?

Valery Forbes:

More details on the statistical methods are needed.

RESPONSE 3.3-VF-5:

Please refer to the response to General Issue 1.C.

Thomas W. La Point:

Yes!

James T. Oris:

The statistical approaches used in the fish section are not always clear and are not fully explained in the document. I spent most of my review time on the ERA reading the fish section, and the more I read, the less clear the statistics section becomes. Thus it is not possible to evaluate whether statistical techniques were properly applied in all cases since insufficient detail is present to make such an evaluation. Analysis of statistical power, when appropriate, is not provided and queries from panelists concerning power analyses were not adequately addressed.

RESPONSE 3.3-JO-11:

Please refer to the response to General Issue 1.C.

The approach described in Appendix C.2 for handling non-detects would appear to be appropriate, but it is not clear that the procedure was used in all cases. Sections in both fish and invertebrate assessments refer to Appendix C.2, but then unless I missed something, non-detects in the invertebrate assessment were set at 0.5 of the detection limit and were set to the value of the detection limit in the fish assessment. In both assessments then, it appears that fixed-replacement was chosen (with little evidence to support the decision), but two different replacements were used. There seems to be a disconnect between the two assessments and it is not clear how these values were chosen. Were the statistical techniques used clearly described, appropriate, and properly applied for the objectives of the analysis?

RESPONSE 3.3-JO-12:

Neither the fish nor benthic invertebrate ERA used a default arbitrary fixed replacement method for estimating non-detected concentrations. Consistent with Appendix C.2, the first step was to conduct a bounding analysis with ND=DL and ND=0 to evaluate the sensitivity to detection limit replacement. In some cases, the analyses were highly insensitive to the choice of replacement method, in which case ND=½DL. EPA agrees that for some of these cases, additional information is required in the ERA to demonstrate the "risk ranges" resulting from the bounding analysis. These will be provided in the revised ERA.

In other cases, such as the calculation of TEQ values in fish tissues, the analysis was more sensitive to non-detected values. In the latter cases, the ERA presented risk estimates made using both bounds to explicitly describe the uncertainty associated with detection limits.

Mary Ann Ottinger:

Representative models, including the range of species and habitats should bracket the sensitivities and exposure scenarios. As such, the models reflect the variety of species and range of habitats. The models should also consider differences in body lipid, which will influence the PCB body burden. More detail is needed to clearly explain the analyses in relation to the range of species and environments that they prefer.

RESPONSE 3.3-MO-4:

Please refer to the response to General Issue 13.C (Lipids in Fish).

Brad Sample:

The statistic analyses appears to be generally suitable and appropriate, however transparency in approach is lacking as discussed in comment 1.e (above).

RESPONSE 3.3-BS-2:

Please refer to the response to General Issue 1.C.

Ralph G. Stahl:

The statistical techniques were clearly described and appropriate.

Timothy Thompson:

No. See previous comments.

3.3(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria?

Valery Forbes:

It is my understanding that some of the deformities observed in the Phase I toxicity study (USGS) are also consistent with Hg and/or PAH toxicity. I did not see this reflected in Appendix F.

RESPONSE 3.3-VF-6:

EPA agrees that other contaminants can elicit adverse responses in fish, including some of the responses observed in the fish toxicity studies. The Reviewers have specifically identified mercury and PAHs as potential causative agents in the EPA studies; however, it is highly unlikely that these contaminants exerted significant responses in the site-specific toxicity studies, for several reasons:

- The processing methods of the tissue extracts used in the Phase II study (described in the December 2003 responses to initial Reviewer questions) would reduce toxicity associated with contaminants other than hydrophobic organic contaminants (e.g., PAH concentrations reduced in processing).

- Mercury was eliminated from consideration as a COC based on comparisons of PSA tissue concentrations to background concentrations. In addition, the concentrations of mercury observed in PSA fish were below ecotoxicology-based benchmarks for effects (see Response 3.3-JO-7 above).

- The route of exposure simulated in Phase II (maternal transfer) is not conducive to the elicitation of PAH effects.

- Offspring in both Phase I and Phase II were reared in a laboratory environment in which PAH concentrations in the water column were low. The exposure paradigms applied by USGS are inconsistent with PAH toxicosis.

EPA provided a receptor-specific screening for fish that excluded pesticides and mercury from further consideration in the ERA. PAHs were retained in the fish ERA and were evaluated via consideration of sediment PAH concentrations.

Please refer also to the response to General Issue 3 (Screening of COPCs and Determination of COCs).

The conclusion of the assessment was ‘low risk’ despite evidence of impairment with respect to the assessment endpoints. Justification (EPA response to Panel Question JO34) is that ‘the magnitude of that harm appears to be sufficiently low as to not result in observed population level effects’. Again this would indicate that it is persistence of fish populations that is the actual assessment endpoint being employed.

RESPONSE 3.3-VF-7:

Please refer to the responses to General Issues 7 and 9.

The bias of field populations toward older individuals should be further considered for other possible explanations than lack of fishing.

RESPONSE 3.3-VF-8:

Please refer to the response to General Issue 13.F.

Thomas W. La Point:

Yes, in part. The risk characterization concluded "A high probability of adverse impacts to fish from tPCBs and/or TEQ…. [but mortality of adults is unlikely]. The magnitude of the risk is high, quite independent of adult mortality. Through the developmental toxicity testing, the literature data, and the field-measured limited recruitment into the bass population, the concluding risk characterization of "low" or "undetermined" magnitude is not, in my opinion, justified. The data are sufficient to determine at least moderate, if not high, magnitude. The fact that fish populations have been measured to be "self-sustaining" could be the result of the limited number of sampling events (= four) over the last 4 to 5 years. It may take longer for the full reproductive effects to show up. Further, the evidence of abnormalities may ultimately be a consideration for the overall population "health."

RESPONSE 3.3-TL-1:

Please refer to the responses to General Issues 13.E and 13.F.

James T. Oris:

The general characterization of risk is moderately supported by the available information and appears to be within reason, but it is my opinion that potential risk to fish is under estimated. This is because of how the assessment endpoint was defined and evaluated, restricted solely to toxicity testing and very high levels of effect being used to derive benchmarks. The simple presence of fish in the river does not mean the population is healthy (as evidenced by the amphibian modeling exercise). The presence of disease and deformities in the PSA indicates further evaluation is necessary. In addition, because of the way HQ values were used, several instances of risk potential were discounted (further explanation is in 3.3(j) below).

RESPONSE 3.3-JO-13:

Please refer to the responses to General Issues 2, 7, 13.E, and 13.F.

Risk terminology used within the fish section needs to be revised and brought into concordance with other assessments (e.g., wildlife sections). Based on current terminology, any HQ value less than 1 presents no potential risk. HQ values greater than one present varying levels of risk on what seems to be a scale of HQ=1-3 (low risk), HQ= 3-10 (moderate risk), and HQ>10 (high risk). However, this scale is neither formally defined nor justified in the document, and the scale seems to be arbitrary. Based on the methods of developing benchmarks (non-conservative) and the use of extremely high-levels of effect, all HQ values greater than 1 should indicate significant potential for risk. Levels of HQs that approach 1 may also indicate risk potential. For example, there is little possibility that there is a difference in risk potential between an HQ=0.9 and HQ=1.1, yet under the current scenario, the HQ=0.9 is considered "no risk".

Application of HQ values in the risk characterization is wholly deterministic, even though a probabilistic treatment could be used (such as in wildlife assessments). It is strongly suggested that the risk terminology and the risk characterization be conducted using the approach used in the wildlife sections.

RESPONSE 3.3-JO-14:

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

Mary Ann Ottinger:

The risk conclusions are supported by data collected and models generated. Reanalysis of the data may result in some increase in risk.

Brad Sample:

The risk characterization as presented in the ERA is probably correct. I do believe that more analyses and discussion of the observation of skewed population demographics (great number of old individuals) should be provided.

RESPONSE 3.3-BS-3:

Please refer to the response to General Issue 13.F.

I was pleased to see the point by point evaluation of fish concentrations Figures F.4-1 through F.4-5, plus F.4-7 and F.4-8.

Ralph G. Stahl:

The characterization of risk is supported by the available information, and was appropriate. During the January 2004 public meetings it was noted that substantial work on the presence of DELTs in local finfish populations had been conducted. This information was not readily evident in the ERA or the appendices and should be discussed more fully in the main text of the ERA.

RESPONSE 3.3-RS-4:

Please refer to the response to General Issue 13.E.

Timothy Thompson:

While I generally disagree with the use and assignment of toxicological and field study data in this assessment, the overall risk characterization is probably valid; While “there is evidence of harm for all measurement endpoints evaluated, but that the magnitude of those risks is generally low” is a qualitative statement, it is likely a valid assessment given the data.

Within the presentations of data for tPAH, I have to agree with the conclusions that risks attributable to PAHs appear low. However, I recommend this issue be re-evaluated.

RESPONSE 3.3-TT-10:

EPA concurs that PAH risks to fish are low. However, because concentrations of potentially toxic metabolic products in fish tissues were not measured, it is not possible to re-evaluate the issue.

Agree that the field surveys support an assessment that there are reproductive populations of large mouth bass within the Housatonic River – and believe that characterizing these results as "lack of catastrophic effects" is a negative disservice to the existing data.

RESPONSE 3.3-TT-11:

Please refer to the response to General Issue 7. EPA agrees that the field surveys indicate that bass are reproducing within the PSA. However, EPA also agrees with other Reviewers who commented on the limitations of the field studies for identification of smaller effect sizes that may be ecologically significant. The assessment endpoint for fish is not simply whether fish reproduce in the PSA, but whether the degree of reproduction is adversely affected, relative to an uncontaminated condition.

I do agree with the statement that the field surveys cannot predict "lesser impacts", but believe those are not well defined in the ERA. The presence of disease and deformities should be put to better use – although it is my understanding that the incidence of these were relatively low.

RESPONSE 3.3-TT-12:

EPA concurs that the "lesser impacts" are not well defined in the ERA. This will be addressed in the revised ERA through the expanded assessment of population demographics (see response to General Issue 13.F) and consideration of disease and deformities (see response to General Issue 13.E).

Finally, the population demographics is puzzling. The statement that the population is dominated by older individuals due to "lack of predation" appears speculative.

RESPONSE 3.3-TT-13:

Please refer to the response to General Issue 13.F. EPA agrees that the largemouth bass population demographics are skewed toward older individuals. EPA has high confidence in the conclusion that predation is constrained in the Housatonic River PSA. This results from the advisory on fish consumption that has been in place for more than 20 years, and also the high risks to piscivorous mammals identified in the ERA. EPA agrees that lack of predation cannot be conclusively identified as the cause of the observed population demographics. Instead, EPA indicated that the observed population demographics are characteristic of an unexploited fish population.

3.3(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made.

Valery Forbes:

On the basis of the site-specific toxicity tests evidence/magnitude of harm to fish (and their associated uncertainties) were presented as hazard quotients, with the variation reflecting only variation in the numerator (i.e., in fish tissue concentrations). Hazard quotients were also used for other receptors (e.g., benthic invertebrates) but I am not entirely sure whether the size of the HQ can/should be interpreted the same for all receptors given that the effects thresholds are based on different kinds of tests.

RESPONSE 3.3-VF-9:

Please refer to the response to General Issue 1.D. EPA concurs that the HQs presented in Section 12 of the ERA reflected variation in the exposure term only. However, the risk characterization considered effects thresholds from multiple lines-of-evidence (literature, Phase I, and Phase II), and the concordance among lines-of-evidence supported the MATC value that was applied. EPA concurs that consideration of multiple effect sizes would improve the risk characterization for fish.

Cumulative plots of tissue burdens that show probability of exceeding thresholds (sent in response to written question from JO) are a useful way to present the data.

RESPONSE 3.3-VF-10:

Please refer to the response to General Issue 1.D. The cumulative plots of tissue concentrations were presented in the ERA document, and were resubmitted in December 2003 based on a Reviewer's request to see the concentration data depicted on an arithmetic scale.

Other COCs with HQs exceeding one should be discussed.

RESPONSE 3.3-VF-11:

Other COCs with HQs above 1 (e.g., PAHs) are discussed in the risk characterization. EPA concurs that these risks should be summarized in the uncertainty assessment, and will include that summary in the revised ERA.

Thomas W. La Point:

They were identified and addressed (particularly the sub-lethal effects in larval and YOY fish), but not appropriately weighted.

RESPONSE 3.3-TL-2:

Please refer to the response to General Issue 8.A. The weights will be reevaluated and more clearly described.

In my opinion, the threshold effects levels should be established lower than they are presently. The effects on larvae and eggs (@ circa 11 mg/kg), determined from literature, should be deemed of sufficient strength to lower the 32 mg/kg threshold for warmwater fish. I would hope there is value in not broaching the "catastrophic" effect level before "high" magnitude is assigned.

RESPONSE 3.3-TL-3:

EPA will review the MATC derivation, based on Reviewer comments. These comments include consideration of disease and deformities (see response to General Issue 13.E), evaluation of population responses (see response to General Issue 13.F), reassessment of toxicity endpoint calculations (see response to General Issue 12), and consideration of multiple effect sizes (see response to General Issue 1.C). EPA concurs that individual toxicity trials (both in literature and site-specific studies) sometimes indicated toxicity at PCB concentrations below 32 mg/kg. These results will be placed into context using the results from all lines of evidence, and considering the consistency of observed responses. Also, EPA assigned greater weight to the site-specific toxicity studies relative to the literature studies because of the lower uncertainty in effects observed to PSA fish rather than other populations.

James T. Oris:

Uncertainties were identified and discussed within reason. Improvements can be made by discussing the uncertainties associated with using median HQ values as deterministic risk values, and discussing the uncertainties associated with multiple toxins present at levels that exceed benchmark values.

RESPONSE 3.3-JO-15:

Please refer to Response 3.3-VF-11 above.

Mary Ann Ottinger:

There were a number of uncertainties, especially associated with the habitat variation. The assumptions appear justified. Additional sampling and data collection may be warranted to address some of these uncertainties and to link data in the lab to observations in the field.

Brad Sample:

The discussion of uncertainties is reasonably adequate. A couple of topics are missing that should be included. These include the need to explicitly discuss how PCB and TEQ concentrations vary seasonally in relation to lipid content of fish and how this may affect temporal and season exposure and potentially risk.

RESPONSE 3.3-BS-4:

Please refer to the response to General Issue 13.C.

In addition fish are simultaneously exposed to multiple contaminants. The interactive effects of exposure to multiple toxins should be discussed.

RESPONSE 3.3-BS-5:

Please refer to Response 3.3-VF-11 above.

Ralph G. Stahl:

The significant uncertainties in the analysis of the assessment endpoints were identified and, for the most part, adequately addressed. Improvement could be made by further discussion of the possibility that COCs other than tPCBs and TEQ could have caused some of the morphological abnormalities observed in the fish. For example, inorganic mercury is known to cause skeletal malformations in finfish yet this is not fully discussed in this section.

RESPONSE 3.3-RS-5:

Inorganic mercury was eliminated from consideration as a COC for fish based on the receptor-specific screening. The rationale for elimination of mercury was that field data indicate that tissue mercury concentrations in Housatonic River fish are similar to background concentrations found in fish collected from the least contaminated areas of MA, and are below concentrations of toxicological concern. Mercury contamination, therefore, is a not a plausible explanation for the increased incidence of morphological abnormalities in Housatonic River fish relative to reference areas. EPA provided additional information on both mercury and PAH in relation to swim bladder abnormalities in the responses to initial review comments (December 2003); this information will be incorporated in the revised ERA.

EPA believes that mercury should be discussed in the receptor-specific screening, and once eliminated, not further discussed in the ERA. The uncertainty assessment should focus on the COCs that remained following the receptor-specific screening.

Please refer also to Response 3.3-VF-11 above.

Timothy Thompson:

The filet conversion and associated lipid variability is not adequately characterized and should be improved in the final ERA.

RESPONSE 3.3-TT-14:

Please refer to the response to General Issue 13.C. EPA will discuss this issue in the uncertainty assessment in the revised ERA.

The issue of a complete representation of the data should also be expanded – particularly as it relates to MATC calculation and risk expression.

RESPONSE 3.3-TT-15:

Please refer to the response to General Issue 13.A. EPA will discuss this issue in the uncertainty assessment in the revised ERA.

The overall use of the field study data should be explored more thoroughly, and any uncertainties presented objectively.

RESPONSE 3.3-TT-16:

EPA believes that the field study data were used appropriately in the ERA in an objective fashion. Reviewers expressed a range of opinions regarding the findings of the field studies, and the EPA evaluation fell within this range.

However, EPA agrees that additional evaluations of field study data would be helpful and will include these in the revised ERA. Please refer also to the responses to General Issues 13.E and 13.F.

3.3(h) Was the weight of evidence analysis appropriate under the evaluation criteria? If not, how could it be improved?

Valery Forbes:

Generally, yes. However, I feel that the field work (p. 5-70) was underweighted. Also the largemouth bass (GE) field study seems to indicate no evidence of harm. EPA claims this study shows undetermined harm. EPA's conclusion either requires further justification or needs to be changed.

RESPONSE 3.3-VF-12:

EPA agrees to reexamine the description of the weighting of the field study data in the revised ERA. Please refer also to the responses to General Issues 8.B, 13.E, and 13.F.

Thomas W. La Point:

No. The literature data and the toxicity exposures should be given more weight.

RESPONSE 3.3-TL-4:

Please refer to Response 3.3-VF-12 and also to Response 3.3-TT-16 above.

From Figures 4-7 and 4-8, I would conclude that a lower threshold of circa 25-30 ng/kg TEQ and circa 30 mg/kg for tPCBs (Figures 4-1 through 4-5.

RESPONSE 3.3-TL-4:

Please refer to Response 3.3-TL-3 above. MATCs will be reevaluated in the revised ERA.

Although it cannot be put "into evidence" for this ecological RA, it must be acknowledged that the conservative thresholds would also be protective of other higher predators in the Housatonic River valley, for example, *Homo sapiens.*

RESPONSE 3.3-TL-5:

Under the terms of the Consent Decree, and consistent with standard practice, a human health risk assessment for the Housatonic River has been performed as a separate assessment.

James T. Oris:

The weight of evidence analysis is not very transparent in the document as presented. Please see comment B.2 above for details.

RESPONSE 3.3-JO-16:

Please refer to the response to General Issue 8.A.

Mary Ann Ottinger:

The WOE is appropriate given the data, interpretations, and models. Reanalysis may change the conclusions, especially for warm water species.

Brad Sample:

The weight-of-evidence (WoE) analyses is generally appropriate but is not adequately transparent.

RESPONSE 3.3-BS-6:

Please refer to the response to General Issue 8.A.

The WoE analyses use best professional judgment to apply final weightings for the various lines of evidence and their attributes. It is unclear how final weightings were derived. Some discussion as to how these weights were derived needs to be provided – were all attributes judged equally? If not need to know why and this needs to be explained – this will allow reviewers to judge the process.

RESPONSE 3.3-BS-7:

Please refer to the response to General Issue 8.A.

Weighting for field studies appears low – rated as undetermined for evidence of harm – need to explain how overall weightings were derived

RESPONSE 3.3-BS-8:

Please refer to the response to General Issue 8.C.

Ralph G. Stahl:

The weight of evidence analysis was appropriate under the evaluation criteria. It could be improved by giving more weight to the field investigations whether conducted by EPA or GE.

RESPONSE 3.3-RS-6:

Please refer to the response to General Issue 8.C.

Timothy Thompson:

See comments above.

Also, the GE field studies appear to be under-rated in the WOE. Recommend that either the weighting be adjusted upward for the field study, or a better discussion needs to be made on why those data are discounted.

RESPONSE 3.3-TT-17:

Please refer to the response to General Issue 8.C.

3.3(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?

Valery Forbes:

The factor of 4 used to extrapolate risk to trout seems somewhat arbitrary. I noted in response to Panel Question JO32 that the EPA responded to the effect that the factor of 4 applied to the fish toxicity was not a safety factor but rather an extrapolation factor to account for potential sensitivity differences. Extrapolation factors and safety factors are actually synonymous (as are application factors and uncertainty factors). In any case justification for the use of the factor 4 in this instance to account for sensitivity differences should be spelled out more clearly in the ERA.

RESPONSE 3.3-VF-13:

Please refer to the response to General Issue 13.B.

Thomas W. La Point:

Yes. The sediment characterizations and concentrations were such to show low risk.

James T. Oris:

Given the data and resources available for these studies, risk estimates for non-site-specific areas appear appropriate, but should carry lower weight in determining risk.

RESPONSE 3.3-JO-17:

EPA agrees that the risk estimates for downstream areas are more uncertain than the estimates for the PSA, and will include text to this effect in the revised ERA.

Further studies downstream of PSA may be needed to confirm what should be considered preliminary conclusions. Further studies should be conducted to determine whether altered population age-structure is due to lack of harvesting of older fish or lack of recruitment of young fish. No in situ fish toxicity tests were done and all data on fish toxicity, thus, is inferred.

RESPONSE 3.3-JO-18:

EPA believes that adequate data exist for the estimation of downstream risks to fish. Any further study of downstream populations would be further complicated by the fish advisories, stocking programs, and activities associated with the fisheries management areas.

Mary Ann Ottinger:

The risk estimates are based on sediment content and on some samples taken from selected sites. As with the benthic invertebrates and the amphibians, there are going to be areas of potential higher exposure, which are not an issue unless the individual encounters those areas. Therefore, the risk decreases downstream, but there remain areas of concern for exposure.

Brad Sample:

The effects data generally appear suitable and appropriate – However, as this is out of my area of expertise, I need to defer to the other panel members on this topic.

Ralph G. Stahl:

In general the risk estimates were objectively and appropriately derived for reaches of the river where site-specific studies were not conducted. However, as mentioned earlier, there is uncertainty and some about the division of the warmwater MATC (49 mg/kg) by 4 for use in assessing potential risks to coldwater species downstream of the PSA.

RESPONSE 3.3-RS-7:

Please refer to the response to General Issue 13.B.

Timothy Thompson:

No. While the methodology is sound overall, the estimation of the MATC should be revisited, based upon the same rules established for benthic infauna.

RESPONSE 3.3-TT-18:

Please refer to the responses to General Issues 6 and 13.A.

Extrapolation to trout should be better explained – and perhaps consider application of uncertainty factors.

RESPONSE 3.3-TT-19:

Please refer to the response to General Issue 13.B.

3.3(j) In the Panel members' opinions, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?

Valery Forbes:

The ERA concluded that risk is low to moderate for fish and that confidence in this conclusion is moderate. I agree that risks of PCBs for the persistence of fish populations are low, but I believe that confidence in this conclusion is high. The extent to which PCBs cause morphological deformities in individual fish is a separate issue, but clearly one of some interest to the public.

RESPONSE 3.3-VF-14:

The Reviewers expressed a range of opinions regarding the importance of individual-level responses (endocrine and histological end points, steroid hormones, gonadal abnormalities, fish disease and deformities, etc.) as indicators of potential population effects. Given this difference in opinion, EPA believes that the best approach is to discuss all measurements (including morphological deformities) in the risk characterization, but to separate these endpoints in the WOE evaluation. This approach will be implemented in the revised ERA.

Thomas W. La Point:

Not completely, for the reasons outlined above.

James T. Oris:

It is my opinion that potential risk to fish in the PSA is underestimated. This is because of how the assessment endpoint was defined and evaluated, restricted solely to toxicity testing and very high levels of effect being used to derive benchmarks. The simple presence of fish in the river does not mean the population is healthy (as evidenced by the amphibian modeling exercise). The presence of disease and deformities in the PSA indicates further evaluation is necessary.

RESPONSE 3.3-JO-19:

EPA will review the MATC derivation and risk characterization, based on Reviewer comments. These comments include consideration of disease and deformities (please see response to General Issue 13.E), evaluation of population responses (please see response to General Issue 13.F), reassessment of toxicity endpoint calculations (please see response to General Issue 13.A), and consideration of multiple effect sizes (please see response to General Issue 1.C).

In addition, because of the way HQ values were used, several instances of risk potential were discounted. As explained in 3.3(f).2 and 3.3(f).3, the use of deterministic HQ values with an absolute cut-off of HQ<1 presents no risk is problematic in the fish assessment. Figures F.4-6 and F.4-9 indicate that the distribution of HQ values for several fish species span the value of 1, even though the median HQ value was less than one. For example, in Figure F.4-6, only three of the listed median HQ values exceed 1. However, six of the HQ distributions have a 20% or greater chance of exceeding the value of 1 (method for defining "low risk" in wildlife sections). Thus, using the criteria set for fish risk characterization versus wildlife risk characterization, one may make quite different conclusions.

RESPONSE 3.3-JO-20:

Please refer to the response to General Issue 7.

Mary Ann Ottinger:

Yes, with consideration of some of the points made in the earlier discussion.

Brad Sample:

Based on the information provided in the ERA, the risk conclusions are probably accurate and appropriate.

Ralph G. Stahl:

In my opinion, the evaluation supports the conclusions regarding low risk to local populations of finfish. The non-lethal, developmental abnormalities are noteworthy and potentially indicative of chronic, subtle effects that may or may not have ecological significance. These same developmental abnormalities do not appear to be reflected in the field observations, nor does the percentage of fish with DELTs appear to be increased above other areas that are not contaminated with tPCBs. Risk may be intermediate or higher for the fish, but the evidence from the field and laboratory studies is simply not concordant. See also (Barnthouse *et al.* 2003)

RESPONSE 3.3-RS-8:

The field studies were not designed to evaluate these specific endpoints. During the EPA fish collections (generally performed at the end of the growing season) observations of gross abnormalities (i.e., DELT data) were made, mostly in adult fish. In contrast, the developmental abnormalities observed in Phase I and Phase II studies were for developing fry (i.e., swim-up stage), and therefore cannot be directly compared to the fish sampled by EPA or GE. Moreover, fish fry exhibiting the major malformations observed in Phase I and Phase II would not be expected to be observed in the field later in the season, since they would not be expected to survive to later life stages.

Timothy Thompson:

As noted above, I generally agree that there are PCB-related risks to fish in the Housatonic River, but that the magnitude of those risks appear to be relatively low. I believe that the data exist within this study to actually represent both probability of exposure AND level of effect (e.g., 85% probability of encountering tPCB concentrations that yield a 20% swim bladder effect). The ERA should be re-expressed in those terms.

RESPONSE 3.3-TT-20:

Please refer to the response to General Issue 1.C.

3.4 Insectivorous Birds

> ***3.4.(a) Were the EPA studies and analyses performed (e.g., field studies, site-specific toxicity studies, comparison of exposure and effects) appropriate under the evaluation criteria, and based on accepted scientific practices?***

Valery Forbes:

Yes, the tree swallow study seems to have been appropriately conducted. However the fact that the tree swallow seems to be relatively insensitive to PCBs causes some concern and perhaps needs to be better argued for in the ERA.

RESPONSE 3.4-VF-1

The tree swallow was selected as a representative species 6 years ago, before much of the current data which suggest that this species is not particularly sensitive to PCBs were available. Please refer to the response to General Issue 14.A.

Thomas W. La Point:

By-and-large, yes. The three-year swallow study incorporated a more appropriate duration to study reproduction in this species. Given the high tissue PCB burdens, the weakly negative correlation between PCB body burden and hatching success (and the lack of clutch size relationship with PCB concentration), the risk conclusions should be given higher weight than that from the robin study.

RESPONSE 3.4-TL-1

EPA concurs with the Reviewer's comment and will note in the revised ERA that the field study results do indicate some evidence of harm. Please refer to the response to General Issue 14.B.

The one -year robin study was not of sufficient duration to quantify any effects on reproduction stemming from PCB exposure and it artificially adds uncertainty to risk estimates. Also, I continue to think that the ERA should give more weight to the swallows, as they are more directly linked to emerging insects from aquatic sediments.

RESPONSE 3.4-TL-2

Generally, scores were higher for the temporal representativeness attribute in the WOE analyses if the field study or samples used in the exposure analyses had been conducted over multiple years. Temporal representativeness, however, is only one of 10 attributes that are considered in weighting a particular line of evidence. Nevertheless, the overall score for the tree swallow field study (i.e., “high” – see Table G.4-5) was higher than for the American robin field study (i.e., “moderate/high” – see Table G.4-6). Also, please refer to Response 3.4-TL-1.

Although robins may be an appropriate surrogate species for bird species with a terrestrial food chain, the demonstrated PCB concentrations in the sediments and floodplain, and limited residues in upland areas, the focus should be on the swallows and other species relying on aquatic insects. In fact, the two species that could be selected to complement each other in terms of insectivory would be swallows and bats.

RESPONSE 3.4-TL-3

Tree swallows and robins were selected as the representative species to allow consideration of both the aquatic and terrestrial insectivorous pathways, knowing that there is a significant amount of contamination in the floodplain soil as well as the sediment.

James T. Oris:

As mentioned by one of the public speakers and discussed in the panel, the conceptual model for insectivorous birds needs to be revised or refined. Lumping robins and swallows into one assessment endpoint is not appropriate since they have different food sources. These should either be separate endpoints or they should be used within this endpoint but treated separately through the final risk characterization.

RESPONSE 3.4-JO-1

Please refer to the response to General Issue 14.C. Robins and swallows will be treated separately in the risk characterization.

I have similar comments on bird assessment endpoints as for fish. The use of individual-based metrics that have been shown to be predictive of higher level effects (e.g., immune function) should be considered and discussed.

RESPONSE 3.4-JO-2

Custer (2002) reported on several metrics from the tree swallow study that may be considered predictive of higher-level effects including EROD induction. However, other data from this study were more directly relevant to survival, growth and reproduction (e.g., hatching success); therefore, the data on biochemical endpoints were given little weight in the weight-of-evidence assessment for insectivorous birds. These data will be more clearly presented and thoroughly discussed in the revised ERA.

Mary Ann Ottinger:

The tree swallow study was done carefully and well designed. Sites of nest boxes and reference sites appear reasonable; however the mean accumulation of tPCBs in tissue residues at the Taconic Valley site was 137 mg/kg [ref G-23], which suggests that these birds may have been affected and not true reference animals. Birds from other reference sites had 8.44mg/kg and 11.7mg/kg, respectively. The study contained basic measures of residues and reproduction. More sensitive end points may have provided some indication of biological response to the COCs that do not approach levels that would impair reproduction. No laboratory studies were conducted on representative species (example: bobwhite quail or mallard ducks) to complement field observations and to confirm risk to ground dwelling birds that have the potential for contact with the sediment. Finally, there were some inconsistencies in the models, which in part originate from the use of measured tissue concentrations in pippers and nestlings as an estimate of maternal deposition. This should have been predictive of impact and suggests the possibility that the highly contaminated individuals may not be the ones to inhabit and nest in that area the next year. Were there any banding, recapture studies done on the tree swallows?

RESPONSE 3.4-MO-1

The Taconic Valley site had elevated estimates of tissue residue concentrations of tPCBs in 15-day old tree swallows based on the results of the microexposure model. The results referred to by the Reviewer are the estimated means from the Monte Carlo analyses (see Table G.2-13). The probability bounds analysis conducted using the microexposure model to estimate tPCB concentrations in 15-day old tree swallows indicated that the Monte Carlo results for Taconic Valley are uncertain (generally, approximately two orders of magnitude between the lower and upper bounds). The uncertainty for the three target sites in the PSA and the other two reference sites is lower.

Even so, EPA concurs with the Reviewer's comment that the Taconic Valley site may not be an appropriate reference site. Although EPA will continue to present the tree swallow field study results from the Taconic Valley site in the revised ERA, this site will no longer be referred to as a reference site.

A number of biochemical endpoints were included in the tree swallow field study including EROD and ECOD induction. These results will be incorporated into the revised ERA.

EPA believes that measured tissue concentrations in eggs and pippers is an accurate indicator of maternal deposition (EPA did not use measured tissue concentrations in nestlings as an indicator of maternal deposition). Given that other contaminant exposure routes for eggs/pippers (e.g., transport from air, rainwater, adult feathers, or nest materials across egg shells and into embryos) are likely to be comparatively much less important, it is reasonable to infer that the PCBs and TEQ congeners measured in eggs and pippers came from maternal deposition.

No mark/recapture studies were done as part of the tree swallow field study conducted for EPA by Custer (2002). Mark/recapture studies are being conducted at a number of locations, including the Housatonic River, at this time by USGS. EPA does not believe that the results of the study will impact the risk characterization for tree swallows. However, if the results of the ongoing study become available before completion of the revised ERA, they will be incorporated into the revised ERA.

Brad Sample:

The EPA studies and analyses were suitable and appropriate. The 3-year swallow study was very comprehensive, used a good experimental design, and produced fairly robust data. Analyses and conclusions appear to be appropriate. The rather limited amount of effects observed to swallows from PCBs is consistent with the lower comparative sensitivity of tree swallows to PCBs than some other birds. This lower sensitivity of swallows should be acknowledged and discussed as results from this species are being used to evaluate effects for other avian insectivores.

RESPONSE 3.4-BS-1

Please refer to the response to General Issue 14.A.

I think that the analyses would have benefited from the evaluation of another species that is more highly exposed to invertebrates in floodplain soils, in particular, the American woodcock. Although the secretive nature of this species would have made collection of site-specific reproduction data problematic, modeled exposure estimates using measured PCB and TEQ concentrations in food would have provided a good indication of exposure and effects.

RESPONSE 3.4-BS-2

EPA believes that the American robin is highly exposed to invertebrates in floodplain soil during the spring and summer (the reproductive period), and therefore is an appropriate representative species for this Assessment Endpoint. As discussed in Appendix G.2.1.4.2 of the ERA, earthworms and litter invertebrates comprise approximately 75% of the robin's diet during spring and summer, which was the timeframe for this assessment.

In addition, EPA believes that the addition of an assessment of risk to American woodcock at this point in the ERA process would result in risk conclusions with high uncertainty due to limited sample sizes because of their life history characteristics (e.g., low nest density with high spatial dispersion).

Therefore, EPA does not propose to add an assessment of risks of tPCBs and TEQ to American woodcock in the revised ERA.

Ralph G. Stahl:

The EPA field study in the tree swallow was appropriate under the evaluation criteria, and based on accepted scientific practices. The modeling of exposure and effects was based on accepted scientific practices except that the exposure estimates for the American robin did not appear to follow methods described in EPA's Wildlife Exposure Factors Handbook.

RESPONSE 3.4-RS-1

The exposure analyses for American robin followed the methods described in Sections 3 and 4 of EPA's Wildlife Exposure Factors Handbook (i.e., decision chart shown in Figure 4-3 of the Handbook). The methods used to estimate exposure for American robin were consistent with those used to estimate total daily intake for other wildlife species in the current ERA (with the exception of the microexposure modeling conducted for tree swallows).

The modeling of effects was reasonable given the lack of specific toxicological data on tree swallows or American robins exposed to tPCBs or TEQ.

Timothy Thompson:

The lines of evidence provided by EPA were appropriate to this site, and are within the accepted practices of ecological risk assessments. Tree swallow field studies have been conducted at several PCB and other contaminated sediment sites; those studies are appropriately cited in the ERA. The mathematical models used to estimate intake and uncertainty are also well-established, and are accepted scientific practice within the risk assessment community. Field sampling procedures, quality assurance project plans, data quality objectives, and supporting study plans for ecological characterization and risk assessment were well articulated, and meet the requirements for data generation under both the RCRA, and the CERCLA authorities.

3.4(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria, based on accepted scientific practices, and incorporated, appropriately in the ERA?

Valery Forbes:

Yes, the American robin study was well conducted and is one of the few studies included in the ERA which provided a formal statistical power analysis (provided in the ET& C 2003 publication).

Thomas W. La Point:

The three-year nest box field study for the swallows were highly appropriate. The use of next boxes (measuring the food intake, contaminant concentrations in diet, and body or tissue burdens) has been shown highly effective in other studies (Kendall, parathion study; Cobb-Hooper metal studies; McMurry, DDT studies). The three-year duration of the study was also approaching "sufficiency," as swallows live six to 8 years.

The one-year robin study appears to have several problems associated with it, not the least of which include the low statistical power stemming from next predation and the one-year duration. The variances provided with this study "cloud the issue" and do not appear to provide much quantitative information useful to the overall ERA for insectivorous birds.

RESPONSE 3.4-TL-4

EPA agrees with the comment, and considered these concerns in the weight-of-evidence assessment conducted in the ERA.

James T. Oris:

There were multiple issues with the GE study that are common to all of the GE studies (e.g., no reference areas, limited study design, too narrow of question asked), but the EPA accounted appropriately for these limitations and incorporated the information appropriately into the ERA.

Mary Ann Ottinger:

The representative species for insectivorous field birds were tree swallows (EPA study) and robins. In the field study in robins, higher egg tPCB was found in the PSA compared to reference sites; however, no differences detected in clutch size with tPCB concentrations. EPA reanalyzed this study, and there were concerns about a small n and inclusion of only active nests in the analysis. It would have been valuable to have historical data and to have some of the individuals banded for recapture studies over several years to determine survivorship of the young fledged in the PSA. Moreover, it would be important to ascertain which individuals are nesting in the PSA, i.e., the same individuals and pairs over successive seasons or an influx of new animals each year. If the latter is the case, then the individuals will not have the same body burden of tPCBs as would animals that repeatedly nest in the PSA.

Brad Sample:

The GE robin study suffers from multiple design and analysis issues. As a consequence, it's conclusions are not strongly supported by the data. First, the authors treat the nests as replicates to test for a treatment effect. In actuality, they are testing a location effect. Whereas all exposed nests were within the 10 year floodplain, all reference nests were over 300 m outside the floodplain. Although PCBs may differentiate between these groups, so might habitat differences (which were not quantitatively measured). Because the 'treatment' groups are spatially distinct and are not interspersed, the study is an example of pseudo-replication (see Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. Ecological Mono. 54:187-211). Sample size is 1 treatment and 1 reference, with multiple subsamples in each. Hypothesis testing analyses are therefore not appropriate. A more suitable analyses would be to perform regression analyses to see if there was an association between PCBs in eggs and nest success parameters.

A potential bias in estimation of robin exposure via eggs exists because sampling was restricted to only viable eggs. Non-viable eggs may have been so due to elevated PCB/TEQ concentrations. Preferential selection of only viable eggs may have resulted in an underestimation of egg-mediated exposure to nestling robins. This potential bias was not discussed in either the GE study nor the ERA. At a minimum, a discussion should be included in the uncertainties section of the ERA.

RESPONSE 3.4-BS-3

EPA agrees with the Reviewer’s comments. This issue was briefly raised in Appendix G.4.5 (Sources of Uncertainty) of the current ERA, and an expanded discussion will be included in the revised ERA.

Another limitation of the study was that it only evaluated data from a single year. Nest success can vary dramatically from year to year. Some analyses of how ambient conditions (i.e., weather and other environmental conditions) for 2001 related to other years would strengthen the argument that uncommon events did not occur that may have biased the data. Similarly, if detailed quantitative habitat data were available for each nest, these could be used to help increase confidence that the results represent a treatment effect (of lack of it) and not a habitat effect.

Ralph G. Stahl:

The American robin study conducted by GE was appropriate under the evaluation criteria. This study was conducted for a single year and the number of nests available for inclusion through the period dropped significantly in the PSA and reference areas due to depredation, abandonment, or other factors not readily identifiable. This resulted in a study with low numbers of nests on which to base conclusions. Nevertheless, the results provided by the limited study were incorporated appropriately into the ERA.

Timothy Thompson:

The line of evidence provided by GE for robin productivity within the PSA are appropriate to this site and are generally within the accepted practices of ecological risk assessment. However, the general lack of documentation concerning the study plan, quality assurance procedures, and DQOs limit the utility of the information provided. The overall study report could be improved by incorporating more specific information from the ARCADIS Scope of Work to GE into the Final Report a discussion of how the areas to be studied were selected, a figure showing all nest sites, and those for which eggs and juveniles were collected for PCB analysis, would improve presentation.

The GE study collected very few eggs and nestlings for tPCB analyses, and does not articulate why so few samples were collected. This factor limits the strength of their argument that there is not a relationship between PCB concentration and nest success.

While the GE studies is one of the few studies included in the ERA which provided a formal statistical power analysis (provided in the ET& C 2003 publication), I would recommend EPA revisit those calculations. My own calculations on the data presented show a very low level of power – not at all consistent with what is presented in the ARCADIS Report.

RESPONSE 3.4-TT-1

EPA concurs with this comment. In developing the ERA, EPA did not have access to the GE scope of work. Thus the information was not available for inclusion at that time; it will be discussed in the revised ERA. In addition, EPA will examine the statistical power analysis of the American robin study in the revised ERA.

In addition to the low power associated with few samples, there is also the unanswered question, why were only “viable eggs” sampled. A key element missing from this study (or at least is not documented) is why weren’t eggs randomly collected immediately after laying. Random collection would have eliminated what appears to be bias: testing only eggs that would hatch. I submit this limits the overall confidence with the robin study, and this uncertainty should be further explored in the final ERA.

RESPONSE 3.4-TT-2

EPA agrees with the comment. Please refer to Response 3.4-BS-3.

A verification by EPA that the analytical portion of the study met the overall guidelines for quality assurance and control set forth in the QA/QC Plan (Weston 2000). The report also would be improved by a discussion of the uncertainties associated with the relatively few number of PCB analyses actually made.

RESPONSE 3.4-TT-3

EPA agrees with this comment. This issue was raised in Table G.4-6 and Appendix G.4.5 (Sources of Uncertainty) and was one of the reasons why the American robin field study was not given a high weighting in the weight-of-evidence assessment.

An additional note is that the ARCADIS SOW to GE stated that in addition to tPCBs, "biota samples may be analyzed for PCB congeners, Appendix 9 pesticides…". The ARCADIS scope is clear that it will depend upon "client preferences" whether these additional analyses were to be conducted. However, the final report does not make that distinction for us. If these analyses were done, then those data should be reported and discussed. If not, a simple note that the study focused solely on tPCBs would be helpful.

A general comment is that the overall ERA is done a disservice by not providing the ARCADIS report as an integrated part of the Appendices. The final ERA should include the GE report CDs as part of the ERA package.

RESPONSE 3.4-TT-4

EPA provided the ARCADIS report to the Reviewers, along with the ERA. This information, along with the EPA Principal Investigator reports, will continue to be available as supplemental information to the revised ERA.

3.4(c) Were the estimates of exposure appropriate under the evaluation criteria, and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?

Valery Forbes:

Yes, in general. The probabilistic exposure models used for insectivorous birds and the other wildlife species are state of the art. Although one could disagree with some of the details of the input parameters, they are all clearly expressed and presented in a nicely organized way.

I question the use of energy content of grasshoppers and crickets to represent emergent aquatic insects for the tree swallow and robin assessments. It would not have been that difficult to make measurements of more appropriate prey species as part of the ERA.

RESPONSE 3.4-VF-2

EPA agrees that measuring the gross energy of site-specific emergent aquatic insects would have reduced uncertainty in the ERA. However, sensitivity analyses indicated that gross energy of prey items is rarely an influential variable in the exposure analyses carried out for wildlife; therefore, developing such information was considered not to be cost-effective. EPA believes the lack of site-specific data for energy content of emergent aquatic insects is unlikely to have significantly influenced exposure estimates for tree swallow or American robin.

Thomas W. La Point:

Yes, for the swallow study. No, for the robin study. Overall, the use of micro-exposure models is laudable! In my opinion, this type of modeling makes the best use of a highly limited data set for contaminant uptake into insectivorous birds. Having said that, it should be recognized that each component of the model (e.g., food intake rate, PCB concentration in prey, maternal transfer) has it's own variance. In my opinion, carrying the variances through the modeling effort should be used to help justify conservatism in risk estimates and not to provide justification for lessening a risk characterization.

RESPONSE 3.4-TL-5

The exposure analyses carried out for tree swallows by EPA involved the use of Monte Carlo analysis and probability bounds analysis to propagate variability and uncertainty in the input variables to the exposure estimates. These exposure estimates were then combined with the available effects metrics to estimate risk. EPA then categorized risk based on the criteria specified in Section 6.7 of the ERA. The criteria used to categorize risk were the same for all wildlife species, except threatened and endangered species, which had more conservative criteria. Where uncertainty was high (e.g., because only one line of evidence was available or because a line of evidence had much uncertainty), the uncertainty was acknowledged in the risk characterization section and assessment conclusions.

The field study for the swallows was of minimally sufficient duration to support estimates of exposure obtained from the literature.

RESPONSE 3.4-TL-6

If the Reviewer was commenting on the comparison of results from the microexposure model to the concentrations measured in 12-day nestling tissues, then EPA believes that the 3-year duration of the field study was sufficient for this comparison. The data used to parameterize the maternal transfer and prey concentration variables in the microexposure model were from the same 3 years and same locations, further validating the comparison.

James T. Oris:

More discussion of other COCs is warranted since it appears that data were actually collected on the COCs but were qualitatively discounted in the ERA (Section G.1).

RESPONSE 3.4-JO-3

EPA agrees that the information in the Custer report should be included in the justification for screening out COPCs, and such information will be included in the revised ERA.

Site specific location of nests and local soil, sediment, and food PCB concentrations would help the exposure estimates.

RESPONSE 3.4-JO-4

Please refer to the response to General Issue 1.

Is it possible to use the TDI model to link site specific data to the exposure of birds at specific nesting sites?

RESPONSE 3.4-JO-5

The TDI modeling for tree swallows was done for each of the sites included in the tree swallow field study and thus estimates exposure of birds at specific nesting sites. The same link cannot be made for the nests included in the American robin field study because GE did not collect earthworm and terrestrial invertebrate samples from the locations used in the field study. The locations of the earthworm and terrestrial invertebrate samples collected by EPA (for another purpose) are outside of the foraging areas of the robins included in the GE field study. For these reasons, EPA will not be able to link the TDI model for American robin to the site-specific nesting locations that were part of the GE field study.

Mary Ann Ottinger:

Estimates of tPCBs rely on the survey information. Tree swallows are known to be relatively insensitive to PCBs; many of the measurement endpoints were also relatively insensitive [see Table 2.6-1 (I.2-47] for listing of endpoints identified as responsive to PCBs. In spite of this, there was significantly reduced hatching success in 1999 and a negative correlation of dead embryos and higher tPCBs (some of the highest ever recorded; 2.7-88) in 1998 and 1999. In the context of refinement of the COCs, the issue of pesticides remains because some of the insects are likely to come from fields in which agricultural pesticides are in use. This is also true for the study on robins, which provided some additional exposure information.

RESPONSE 3.4-MO-2

EPA agrees that the information in the Custer report should be included in the justification for screening out COPCs, and such information will be included in the revised ERA.

An additional issue is that of the relative diversity of species found in the areas of continued contamination. More information on ducks would have been helpful, especially in the context of providing a broader view on the chemicals that bioaccumulate in avian species that are constantly in contact with the aquatic environment.

RESPONSE 3.4-MO-3

Based upon the suggestion of the Reviewers, EPA collected and analyzed samples of wood duck eggs from the PSA in May 2004. The results and discussion of potential risks to waterfowl will be included in the revised ERA.

Brad Sample:

Although the estimation of exposure and effects was generally appropriate, there were several areas that require further evaluation:

- It was good to see the development and application of the microexposure model for swallow nestlings. It provides a useful mechanism to evaluate exposure and potential effects for what is clearly the most sensitive life-stage. However, I think that the TDI model should also have been run for adult swallows. This is for two reasons: 1) exposure and risks for adults differ from what would be expected for nestlings; and 2) the TDI modeling approach was used for all other wildlife receptors. To be consistent with, and allow comparisons to other wildlife receptors, the TDI model should also be run for adult swallows.

RESPONSE 3.4-BS-4

EPA did estimate total daily intake for adult tree swallows (see ERA Section G.2.1.1).

- Appendix G, Page G-16, Lines 17-27. It appears that the food ingestion rate for nestling swallows was estimated using the allometric model for adult birds from Nagy (1999). The uncertainty or appropriateness of this model to represent nestling food ingestion should be discussed. Nichols et al. (1995) developed a model for nestling food ingestion in their swallow model. At a minimum, food ingestion rates estimated by both approaches should be compared.

RESPONSE 3.4-BS-5

EPA agrees with the comment and will review the Nichols et al. (1995) model for nestling food ingestion rate and will compare the Nichols model with the allometric model, and a discussion will be added to the revised ERA.

- The TDI modeling effort for robins relied exclusively on directly measured concentrations of PCBs and TEQ in earthworms. Similarly, the microexposure model relied exclusively on PCB and TEQ concentrations in stomach contents. This approach will give good estimates for all locations from which these site-specific biota data were available. However it does not address exposure and risk in areas from which site-specific biota data were unavailable, nor does it allow exposure and risk levels to be associated with specific PCB and TEQ concentrations in floodplain soils or sediments (information extremely useful to make remedial decisions). No effort was made to model bioaccumulation into biota and estimate exposure at all locations, with suitable, species-specific habitat, from which abiotic data were available. Given the large size of the site, the heterogeneous nature of contamination, and the variable movement patterns of resident wildlife, the current approach may not adequately represent the exposure that local populations of wildlife may experience at the site. It is recognized that modeled concentrations in biota increase uncertainty. However, this would be balanced by the increased spatial resolution in exposure estimates. I do not recommend replacing the current approach. Rather, exposure based on COPC concentrations modeled in biota should be used to provide a spatially more complete estimate. The results from both can be compared as part of the risk characterization.

RESPONSE 3.4-BS-6

EPA agrees with the comment that more spatially complete exposure estimates would reduce uncertainty of exposure for some wildlife species (e.g., American robin, short-tailed shrew). In the revised ERA, EPA will either: (1) compare soil concentrations at locations where prey data were available (e.g., earthworm and terrestrial invertebrate sampling locations) to soil concentrations from other areas of the floodplain, or (2) develop a BSAF model to enable exposure estimates across the PSA. Either approach should provide a better characterization of risk across the entire PSA for those wildlife species that had risk estimates at only a few selected locations.

- The TDI and microexposure modeling efforts relied exclusively on directly measured concentrations of COPCs in dietary items (either earthworms for robins, or stomach contents for swallows). The results of these analyses, derived from data representing a comparatively limited spatial area, were then extrapolated to represent the whole PSA. For this approach to be valid, it is important that the areas from which biota data were collected represent the range of possible COPC concentrations in abiotic media, in particular it is essential that the highest concentrations be represented. It is not apparent that any analyses were conducted to indicate that the available site-specific biota adequately represented the full range of possible COPC exposures (including maximum concentrations).

RESPONSE 3.4-BS-7

EPA agrees with this comment. No data on concentrations in earthworms or stomach contents for swallows are available beyond what was used in the ERA. As a result, a surrogate approach is required to respond to this comment. EPA will estimate median sediment concentrations at each of the locations where exposure modeling was done (the median of samples within foraging range of species of interest will be calculated). These medians will then be compared to medians for other areas in the PSA with suitable habitat. The comparison should reveal whether the exposure modeling locations represent the full range of possible COPC exposures.

- References for sources of parameter values in Tables G.2-2 and G.2-3 and Tables G.2-24 and G.2-25 are lacking and should be provided.

RESPONSE 3.4-BS-8

EPA will add the sources to the appropriate tables in the revised ERA.

Ralph G. Stahl:

The estimates of exposure may not be appropriate under the evaluation criteria due to the extensive use of modeling, and thereby the insertion of subjectivity into the assessment. In itself, the modeling work for exposure was reasonable. Site-specific information was available for the dietary intake of prey items by tree swallows, and EPA's Exposure Factors Handbook should be consulted to estimate exposure for American robins.

RESPONSE 3.4-RS-2

As was done for all wildlife exposure analyses, the Wildlife Exposures Handbook (EPA, 1993) was consulted to estimate exposure for American robins. The total daily intake models used for American robin and other wildlife species are those described and recommended by the Wildlife Exposures Handbook. For most input distributions, EPA consulted this Handbook and then located the primary references listed in the Handbook to ensure that the results listed in the Handbook were correct. In all cases where primary references were located and reviewed, they were cited in the ERA.

Timothy Thompson:

No. There is a discontinuity between the screening procedure applied in Appendix G, and the field study reports that should be bridged. Section G.1 states that the other COPCs, and particular the organochlorine pesticides, are screened out "because their actual concentrations in the PSA were likely much lower than the measured values due to laboratory interference". However, the Custer report includes analyses and detection of polycyclic aromatic hydrocarbons (no detections) and aliphatic hydrocarbons, metals, and several organochlorine compounds including DDT and derivatives. That appendix presents a credible argument for why these other detected COPCs are not at levels of environmental concern in tree swallow tissues. The ERA should be amended to incorporate those findings as a reason for refining the COC list for the probabilistic risk assessment.

RESPONSE 3.4-TT-5

EPA agrees that the information in the Custer report should be included in the justification for screening out COPCs and will include this information in the revised ERA.

3.4(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria?

Valery Forbes:

Yes, the combination of literature-based effects thresholds and field study results was appropriate.

Thomas W. La Point:

Yes. Focusing on reproductive consequences and early life-stage consequences of PCB exposure is highly appropriate. Unfortunately, the study duration to quantify the effects needs to match the life history of the species under consideration. As has been quantified with exposure to other poly-chlorinated hydrocarbons, the "ultimate" effect may be observed in the offspring of nestlings exposed to high PCB levels.

James T. Oris:

Given the statement of the assessment endpoint, the metrics were appropriate. However, there needs to be a discussion of why other endpoints were not used that also would be appropriate. It could be argued that these other endpoints are "individual" level, but many of these endpoints (e.g., physiologically based: gonad weight, enzyme levels, hormone assays, immune function) have been shown to be predictive of long-term effects in birds.

RESPONSE 3.4-JO-6

Custer (2002) reported on several metrics from the tree swallow study that may be considered predictive of higher-level effects, including EROD induction. However, other data from this study were more directly relevant to survival, growth, and reproduction (e.g., hatching success); therefore, the data on biochemical endpoints were given little weight in the weight-of-evidence assessment for insectivorous birds. These data will be more clearly presented and thoroughly discussed in the revised ERA.

Mary Ann Ottinger:

Some more targeted endpoints should have been measured to detect effects associated with moderate stress or even a hormesis response. This would have provided a more complete overview of the organisms' responses and level of stress in the environment. Furthermore, as the area has been contaminated for a number of generations, it is possible that the population now resident in the PSA is a selected subpopulation that has higher resistance to the COCs.

The reproductive axis often continues to function despite environmental challenges up to some point, at which time the system is likely to collapse. This collapse, often associated with heightened adrenal system activity then results in the cessation of egg production and testicular regression. Tree swallows appear to be very hardy and are able to persist in reproduction in spite of environmental challenges. They may however, have altered physiological responses (ex. EROD 2.7-71) and associated alteration in biochemical measures from the PCB exposure that would provide a reliable index of exposure, be indicative of lower level exposures, and be predictive of some level of adverse impact. These endpoints could include steroid hormones and hypothalamic neurotransmitters, which would provide potentially more sensitive endpoints.

RESPONSE 3.4-MO-4

Please refer to Response 3.4-JO-2.

Brad Sample:

The effects data appear to be suitable and appropriate. EPA has conducted a very thorough review of the available toxicity data and has made a very strong effort to adequately represent the range of potential effects. One aspect lacking from the ERA was an evaluation of egg-based effects data. Site-specific egg-based exposure data were available for robins (and could potentially have been modeled for swallows). Egg-based PCB effects data are available in the literature (Hoffman et al. 1996, "PCBs and dioxins in birds." In *Environmental Contaminants in*

Wildlife, Interpreting Tissue Concentrations. W. Nelson Beyer, Gary H. Heinz, and Amy W. Redmon-Norwood, eds. Lewis Publishers, Boca Raton.).

RESPONSE 3.4-BS-9

EPA concurs with this comment and will consult the suggested reference and compare site-specific measured concentrations in eggs (where available) to relevant effects metrics from the literature in the revised ERA.

Ralph G. Stahl:

The effects metrics were identified and appropriate under the evaluation criteria. Both the EPA and GE studies evaluated effects on reproductive success of the tree swallow and the American robin. These metrics are difficult to evaluate but were done so in a reasonable and objective manner.

Timothy Thompson:

Yes. No further comment here.

3.4(e) Were the statistical techniques used clearly described, appropriate, and properly applied for the objectives of the analysis?

Valery Forbes:

Yes, well done.

Thomas W. La Point:

The use of microexposure and daily intake modeling and the risk estimates from upper and lower risk bounds (and Monte Carlo simulations) are excellent approaches.

James T. Oris:

The TDI model was appropriate and well-described. Statistical techniques need to be described in more detail so that it is clear how the data were analyzed throughout the document.

Mary Ann Ottinger:

The models generated were complex due to the complexity of the ERA [see 2.6-26 for risk curves relative to low, intermediate, and high risk categories]. This approach is sound and has been conducted in a most thorough manner. The criteria for estimating risk are as follows: representative species criteria for risk=20% probability for exceeding threshold for most sensitive species=low risk; greater than 20% probability for exceeding the threshold for the most tolerant species=high risk. In some cases, this yields a large intermediate range and especially in the case of the tree swallow may artificially inflate the intermediate range and ultimately underestimate risk to other passerines. The integration of the two models (TEQ and microexposure) is not transparent and it is not clear which (or if both) model has been used for the summarized ERA. A more detailed and clearer explanation should be included.

RESPONSE 3.4-MO-5

EPA concurs with the comment that use of threshold ranges sometimes created large intermediate risk ranges. However, given the limitations of the toxicity data used to develop the threshold ranges, relatively few options are available to deal with this issue. Therefore, no changes will be made to the risk criteria in the revised ERA.

EPA assumes that the Reviewer meant integration of the TDI (rather than TEQ) and microexposure models for tree swallows. EPA used both approaches to estimate exposure of tree swallows to tPCBs and TEQ. The TDI modeling was done to facilitate comparison of exposure estimates to other wildlife species, but was not used to estimate risk. The microexposure modeling results were used to estimate risks of tPCBs and TEQ to tree swallows. The two models were not "integrated" in any way. A better explanation of how the models were used in characterizing risk to tree swallows will be provided in the revised ERA.

The TDI exposure models for tree swallows are reasonable and reflect the higher residues found at the Taconic Valley site. The TDI exposure model for TEQ consequently shows a very wide intermediate range with a relatively high dose to achieve the intermediate-high criterion. This reflects the relatively insensitivity of the tree swallow to PCBs and is likely to underestimate risk to other insectivorous species. Other species, including the ground dwelling galliformes, pheasants, ducks, and geese are likely to be more sensitive to PCB impact. Data from these species would have provided a more inclusive and representative data set and a more accurate Monte Carlo analysis. In spite of this shortcoming, the models consistently predicted a higher risk than was found in the field studies, suggesting that there are uncertainties in the data set (in part due to the choice of representative species among other factors). Finally, a broader range of insectivorous birds should have been sampled and field studies should have included a subset of these species.

RESPONSE 3.4-MO-6

Based upon the suggestion of the Reviewers, EPA collected and analyzed samples of wood duck eggs from the PSA in May 2004. The results and discussion of potential risks to waterfowl will be included in the revised ERA.

Brad Sample:

Similar issue with transparency of statistical analyses as discussed for earlier sections, apply to this section too.

RESPONSE 3.4-BS-10

Please refer to the response to General Issue 1.

It is not immediately clear in this section, nor in the subsequent sections and appendices in which Monte Carlo simulation of wildlife exposure is discussed, as to how chemical concentration data were included in the exposure models. However, at the December meeting, Dwayne Moore stated that chemical concentrations were included in the Monte Carlo analyses as a point value (the 95% upper confidence limit on the arithmetic mean [UCL95]). I am concerned that this approach does not adequately nor accurately represent the variation and uncertainty in exposure. The exposure distribution that is generated only represents the uncertainty and variability about the distributions for the life-history based parameters (body weight, diet composition, food ingestion, etc.). Variability and uncertainty associated with PCB/TEQ concentrations are not represented. Depending on how large the sample size is and the variance associated with the data, use of the UCL95 may over represent exposure. Although the probability bounds analyses captures what should be all possible exposures, it is a very broad brush. The Monte Carlo analyses should provide what is believed to be the reasonable, most-likely exposure distribution based on expected distributions of all parameters.

RESPONSE 3.4-BS-11

Please refer to the response to General Issue 10.

Statistical analyses should be provided for data presented in Tables G.19 through G.23 to indicate which sites differed from each other.

RESPONSE 3.4-BS-12

EPA agrees with this comment and will perform the suggested analyses and provide this information in the revised ERA.

Ralph G. Stahl:

The statistical techniques were clearly described and appropriate.

Timothy Thompson:

Yes, with two exceptions noted below.

The ERA reports that there is a small, but statistically significant negative correlation between tPCBs and hatching success in tree swallows. This is noted throughout the document including the Executive Summary, is noted in the Quantitative Measure attribute, and the Risk Summary. However, it is somewhat misleading, and is somewhat incongruous inconsistent with the analysis presented by Dr. Custer in her report. To quote:

> “There was a negative relationship between total PCBs and hatching success in 1998 and 1999 combined (P = 00.049, Fig.4), **but not when 2000 was added to the model (P=0.273)** [emphasis added].

The fact that 1998 and 1999 were significant when combined is very important – and better use can be made of the data. The statistical relationship when the 2000 data is added is tenuous at best, and I would not go as far as to even label it as a “weak” relationship.

RESPONSE 3.4-TT-6

Please refer to the response to General Issue 14.B.

While I agree with the ERA's position that the statistical analysis conducted in the ARCADIS report were likely not the appropriate choice given the limited data size. However, if the ERA is going to find "fault" with the statistical tests employed, those need to be articulated. In my opinion, given that the sample sizes are small, and the test between the PSA and reference nests are unequal sample size, the more appropriate test to use would have been the Kruskal-Wallis test for non-parametric comparison, but I suspect given the very low difference between means, this too would show now difference.

RESPONSE 3.4-TT-7

In the revised ERA, EPA will expand the discussion of the appropriateness of statistical tests employed in the ARCADIS report.

As noted, I cannot repeat the optimistic power from the ARCADIS report data, and have recommended that EPA look at those data themselves. I would further submit that a power analysis of the tree swallow study would also be required for the reproductive metrics only.

RESPONSE 3.4-TT-8

EPA agrees with this comment, and, in the revised ERA, will present power analyses for both the American robin and tree swallow studies.

3.4(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria?

Valery Forbes:

For the wildlife in general the risk characterizations were the most quantitative. For example, defining risk in terms of 'the probability of a certain percent effect occurring' clearly has advantages over hazard quotients.

Thomas W. La Point:

For the swallows, the characterization of risk, using a WOE approach, appears to be low. That is, the findings included high body burdens, weak – but statistically significant – negative correlation with hatching success, and measured uptake from food items. The only evidence of "no effect" comes from the lack of success on clutch size, a life history characteristic closely controlled by genetic factors. Given this evidence, and the limited value of the robin study, a ranking of "low risk" for tree swallows is questionable, in my opinion.

RESPONSE 3.4-TL-7

Please refer to the response to General Issue 14.B.

James T. Oris:

Yes, appropriate.

Mary Ann Ottinger:

The risk characterization was supported by the data; however, as discussed above, there were some issues associated with the selection of species and in the outcomes. Moreover, the models resulted in varying predictions due to some of the data used to generate the models, with the microexposure models appearing to be more predictive of a range of species. Clarification and possible reanalysis is recommended, with consideration of aligning observed effects with the models generated.

RESPONSE 3.4-MO-7

EPA concurs with the Reviewer's comment that the microexposure models produced reasonable estimates of exposure for tree swallows (see discussion in Appendix G, Section G.2.1.3.5). This model was used in estimating risks to tree swallows only, because the goal was to estimate exposure for growing juvenile nestlings for which corresponding field effects data were available. For other wildlife species, the focus was on estimating exposure to adult females (to correspond to available effects data) and thus microexposure modeling was not required. At this time, EPA does not propose to conduct microexposure modeling for species other than tree swallow.

Brad Sample:

Although the risk characterization for insectivorous birds was generally supported by the available information and was appropriate, multiple issues were identified that require attention/revision:

- The purpose of the risk characterization stated at the end of Section G.4 is to determine whether exposure is 'sufficient to cause adverse effects to individuals inhabiting the PSA'. This purpose should relate back to the assessment endpoint described in the Problem Formulation – I believe that the assessment is supposed to be on populations, not individuals. In addition, some discussion of the spatial aspects of exposure and risk should be included. This is a particularly important issue with a large site such as the Housatonic.

RESPONSE 3.4-BS-13

Refer to the response to General Issue 9.

- To interpret the significance and general applicability of the risk characterization results to the PSA as a whole, it is important to know how the limited number of locations represented by field data or exposure estimates compare to the overall contamination within the PSA as a whole. If the maximum exposure is not represented by these data, then the full range of exposure (and risk) is not adequately represented.

RESPONSE 3.4-BS-14

EPA agrees with the comment that more spatially complete exposure estimates would reduce uncertainty of exposure for some wildlife species (e.g., American robin, short-tailed shrew). In the revised ERA, EPA will either: (1) compare soil concentrations at locations where prey data are available (e.g., earthworm and terrestrial invertebrate sampling locations) to soil concentrations from other areas of the floodplain, or (2) develop a BSAF model to enable exposure estimates across the PSA. Either approach should provide a better characterization of risk across the entire PSA for those wildlife species that had risk estimates at only a few selected locations.

- Because swallows and robins represent very different exposure regimes, I do not think that they are directly comparable and would recommend that they be separated and evaluated independently.

RESPONSE 3.4-BS-15

These species will be evaluated independently in the revised ERA. Please refer to the response to General Issue 14.C.

- The discussion in Section G.4.1.1 is based solely on the evaluation of exposure and risks to nestling swallows. This should be made clear. Because a TDI model was not run for adult swallows, exposure and risks to adults have not been directly evaluated. The text in this section discusses tree swallows in general although data are only for nestling swallows. Additional discussion is needed to explain how the results for nestlings relate to the larger swallow population as a whole.

RESPONSE 3.4-BS-16

TDI modeling was conducted for adult tree swallows, as described in Appendix G, Section 2. However, only the results of the microexposure modeling for juveniles were used to estimate risks. EPA will clarify in the revised ERA that estimated and measured tissue concentrations were used to estimate risks to juvenile tree swallows. The significance of these results with regard to the population of tree swallows will also be discussed. Text will be added to Section G.4.1.2 to clarify that the exposure and risk estimates for robins are based on dietary exposure, not tissue concentrations as was done for swallows.

It should also be noted that risks are estimated based on measured and estimated tissue concentrations – this is different than for robins (next section).

RESPONSE 3.4-BS-17

EPA will clarify this point in the revised ERA.

- The first sentence of the conclusions portion of Section G.4.2 states that fecundity of swallows does not appear to be affected by contaminants. This directly contradicts the results presented in the preceding section, in which a significant negative relationship between PCBs in eggs and hatching success was observed. This discrepancy needs to be resolved. It might be more appropriate to state that swallows are affected but not severely so.

RESPONSE 3.4-BS-18

EPA agrees that the wording of the conclusions could be improved and will do so in the revised ERA. Please refer to the response to General Issue 14.B.

- Analyses of variance methods are used to evaluate whether PCBs had a significant effect on nesting success of robins. Use of regression methods, similar to those used for the swallows would probably be more suitable.

RESPONSE 3.4-BS-19

While EPA concurs with this comment, relatively few egg samples (n=2 for reference location nests, n=9 for target location nests) were analyzed in the robin study. The low number of samples precludes the regression-based approach suggested by the Reviewer.

Ralph G. Stahl:

The characterization of risk was supported by the available information and was appropriate under the evaluation criteria.

Timothy Thompson:

Based on the evidence presented, the ERA in my opinion correctly assigns a "low risk" from exposure to tPCBs and TEQs to insectivorous birds. While having some general concerns with the assignment of "weights" to the metrics, the ERA correctly relied on the field study data to ultimately assess risk.

3.4(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made.

Valery Forbes:

Yes. Possibly with the exception of using a point estimate for concentration of PCBs in prey. Tree swallow HQs were not based on site specific data that could have been used, but rather on the basis of modelled data. I would recommend using appropriate site-specific data whenever possible to reduce uncertainty.

RESPONSE 3.4-VF-3

The modeling effort for tree swallows used site-specific samples (samples taken directly from nestling tree swallows) to parameterize input distributions for dietary concentrations. The upper bound for the effects threshold metric was also from the site-specific field study by Custer (2002).

Thomas W. La Point:

The uncertainties were very well taken into account. However, for these species, I would recommend that uncertainty add to a higher risk estimate, as the species are relatively long-lived and were only sampled over a limited period of their lives.

RESPONSE 3.4-TL-8

The exposure analyses carried out for tree swallows by EPA involved the use of Monte Carlo analysis and probability bounds analysis to propagate variability and uncertainty in the input variables to the exposure estimates. These exposure estimates were then combined with the available effects metrics to estimate risk. EPA then categorized risk based on the criteria specified in Section 6.7 of the ERA. The criteria used to categorize risk were the same for all wildlife species, except threatened and endangered species, which had more conservative criteria. Where uncertainty was high (e.g., because only one line of evidence was available or because a line of evidence had much uncertainty), the uncertainty was acknowledged in the risk characterization section and assessment conclusions.

EPA believes that this process is an objective means of categorizing risk. Upwardly adjusting risk estimates because species are long-lived or for other reasons introduces bias to the process of categorizing risks. The practice would also be very subjective. EPA does not propose to alter the risk categorization process for tree swallows or other wildlife species in the revised ERA.

James T. Oris:

Uncertainties were appropriately described.

Mary Ann Ottinger:

There were a number of uncertainties, already discussed above. As such, there are issues that must be addressed in order to draw realistic conclusions regarding the impact to insectivorous birds. This may require additional field surveys, including egg collections and analyses for PCBs and for other environmental chemicals. As in the frogs, analysis of thyroid hormone is an important endpoint that was ignored in all these studies, yet the relationship to PCB exposure has been established. Alterations in the thyroid hormone axis would be predicted to influence timing of nesting, as well as viability and hatching of offspring. Further, in the robin study, growth rate was slightly reduced (G59), similar to findings in chickens. In the study in chickens, longer exposure also resulted in reduced fertility and productivity. This suggests that the study should have been carried out longer and with more birds. Also, the number of nests should be increased and better documentation of abandoned nests should be conducted. There should be bird survey information available over several years in order to discern possible differences in species richness and nesting densities in the PSA and reference areas.

RESPONSE 3.4-MO-8

EPA agrees that the suggested research initiatives would add to understanding of risks posed by tPCBs and TEQ to insectivorous birds. However, EPA believes that the current information used in the ERA is adequate to evaluate risk to insectivorous birds.

Brad Sample:

Although the uncertainties are generally well captured in the ERA report, the relative importance of various uncertainties is not discussed. Interpretation and understanding of the results would benefit by such an analysis. Spatial heterogeneity of contamination and the degree to which the exposure data actually represent the PSA are very significant issues.

These need to be discussed.

RESPONSE 3.4-BS-20

EPA agrees with this comment, and a discussion of the relative importance of the sources of uncertainty will be provided in the revised ERA.

Ralph G. Stahl:

The significant uncertainties in the analysis of the assessment endpoints were identified and adequately addressed.

Timothy Thompson:

An important source of uncertainty not discussed is the measurement error that likely occurred in PCB measurements between the two field studies. Data variability is a consistent issue throughout the entire ERA and RCRA Facilities Investigation. Measurement error is cited in Section 7.3 of the Work Plan (Weston 2000) as a source of uncertainty. This uncertainty is not specifically called out or accounted for in Appendix G.

RESPONSE 3.4-TT-9

EPA conducted an extensive analysis of both inter- and intralaboratory measurement error, and reported the results in Appendix C.11 to the ERA. EPA agrees that aspects of this information should be incorporated into the main text of the ERA and will include a discussion of the uncertainty associated with measurement error in the revised ERA.

An evaluation of data quality and interpretation issues is presented in Appendix D to the RFI. While that is not specific to bird tissues, the fact that the RPDs in fish tissue duplicates and splits is reported to be approximately 29% for intra-lab comparisons, and 41% for inter-lab comparisons would likely also apply to analyses of invertebrate (prey), egg, and nestling tissues. While this variability cannot be controlled at this stage of the ERA, accounting for it in the uncertainty section is important – recognizing that it will most likely be a qualitative discussion. Having said that, noting that EPA and GE have split samples for several media, I would be surprised if this was not done for these studies, as well. If that is the case, then lab variability must be accounted for explicitly.

RESPONSE 3.4-TT-10

Please see Response 3.4-TT-9. The evaluation of data quality included in Appendix D to the RFI was based, in part, on the analysis conducted by EPA and reported in Appendix C.11 to the ERA.

3.4(h) Was the weight of evidence analysis appropriate under the evaluation criteria? If not, how could it be improved?

Valery Forbes:

The total value of each measurement endpoint is described as a weighted average of the 10 attributes going into the WOE (p. G 77). Correction needed. This was not a weighted average. I believe that earlier chapters in the ERA did not describe the process by which the total values were determined. This should be remedied.

RESPONSE 3.4-VF-4

Please refer to the response to General Issue 8.

My assessment of the American Robin field study (at least the version published in ET&C 2003) was more positive than EPA's and I would have given it somewhat more weight in the assessment than was done.

RESPONSE 3.4-VF-5

The American robin field study was given a moderate/high weighting in the weight-of-evidence assessment for robins (see Table G.4-6). EPA also believes that the study provided useful information to characterize and understand risks of tPCBs and TEQ to robins. However, the study had several shortcomings (e.g., small sample sizes to characterize exposure to embryos and nestlings, sampling of only viable eggs from nests, field study covered only one reproductive season) that prevented it from having a high weighting in the weight-of-evidence assessment. Please refer also to the response to General Issue 8.

Once again, I find the WOE scoring sheets can be ambiguous. A combination of 'no evidence' for a 'low magnitude' of harm (i.e., both field studies) could either be interpreted that the study would have been able to detect a low magnitude of harm but didn't or that the study did not find evidence for harm but there nevertheless could have been low magnitude harm that was not picked up. Reconsideration should be given to the table inputs so that they can stand alone and be interpreted without ambiguity.

RESPONSE 3.4-VF-6

Please refer to the response to General Issue 8.

Thomas W. La Point:

In my opinion, the WOE was not appropriate for these species. I might suggest using the variances associated with feeding, uptake, and maternal transfer and use the variance estimates to indicate the degree of uncertainty (it is acknowledged to be high in this ERA). As such, the WOE approach may not be as useful to estimate long-term results, stemming from measures taken on an "acute" (one year for the robins) or "limited" (three years for the swallows) basis. The only way to improve the analysis is to re-do the robin study, for a minimum of two or three years, perhaps including banding adults and their offspring. Also, if a limited number of swallows were to be banded and studied over three-to-four years, it would help measure effects on the reproductive success of the F1 generation.

RESPONSE 3.4-TL-9

Variability and uncertainty associated with the input variables used to estimate food intake rate (e.g., field metabolic rate, body weight, assimilation efficiency and gross energy of prey items) and maternal transfer were explicitly incorporated in the exposure models conducted for tree swallows and American robins (see Appendix G, Section 2). The WOE assessment used this modeling approach as one line of evidence and explicitly defined probabilistic criteria for determining whether the receptors were at low, intermediate, or high risk. Thus, EPA maintains that such an application of the WOE approach is appropriate for insectivorous bird species.

Due to practical limitations, it is not possible to repeat the robin study or carry out a banding study for tree swallows in support of the ERA. EPA believes that the information used in the ERA is sufficient to characterize risk to insectivorous birds. Please see also related Response 3.4-MO-1.

James T. Oris:

See previous comments on WOE analysis. In addition, the EPA should consider down-weighting the Robin study due to its limitations and deficiencies.

RESPONSE 3.4-JO-7

Please refer to the response to General Issue 8 and Response 3.4-VF-5.

Mary Ann Ottinger:

The WOE is appropriate under the findings from the representative species. As pointed out above, there are difficulties with the representative species chosen and with some of the endpoints that were not included in the assessments. Similarly, the study with robins had experimental flaws due to difficulties in the field study. Consequently, the conclusions drawn using a WOE approach are not compelling in this receptor.

Brad Sample:

As stated previously, the weight-of-evidence (WoE) analyses is not adequately transparent. The WoE analyses use best professional judgment to apply final weightings for the various lines of evidence and their attributes. It is unclear how final weightings were derived. Some discussion as to how these weights were derived needs to be provided – were all attributes judged equally? If not need to know why and this needs to be explained – this will allow reviewers to judge the process.

RESPONSE 3.4-BS-21

Please refer to the response to General Issue 8.

The WoE analyses for swallows appears to be appropriate and suited to the available data, except that the evidence of harm from the field study is incorrectly stated as 'No'. This is not consistent with earlier text. On page G-78, Line 15 the statement is made that risks are present but low for swallows. Consequently, I think the table should be revised to state 'Yes' for evidence for harm based on the field study.

RESPONSE 3.4-BS-22

EPA agrees with the Reviewer's comment, and will correct the table in the revised ERA.

The WoE analyses for robins generally appears to appropriate and suited to the available data, except that I believe that the field study is weighted to high. I think that it should have a weighting no higher than moderate.

RESPONSE 3.4-BS-23

EPA will revisit the WOE for this and other studies to ensure that they are consistently evaluated. Please see response to General Issue 8.

Ralph G. Stahl:

The weight of evidence analysis was not appropriate under the evaluation criteria. The modeled exposure and effects was given a weight of "moderate" which does not appear to be objective nor reasonable. Given that there were two field studies, one for tree swallows and one for American robins, the modeled exposure and effects should be given a "low" weighting in the analysis.

RESPONSE 3.4-RS-3

In the WOE approach, each line of evidence is evaluated and weighted independent of whether other lines of evidence are available. Tables G.4-5 and G.4-6 provide a justification for why the modeling line of evidence was given a moderate weighting for tree swallows and American robin, respectively. Also, please refer to the response to General Issue 8.

Timothy Thompson:

For insectivorous birds, there are several instances where the weight assignments do not appear to meet the criteria of "objectivity, consistency, and reasonableness".

For the analysis for stressor/response, the ERA assigns the EPA study a "moderate/high" value, but assigns the GE study only a "moderate" value. The principal reason, as worded in the text, is that the study did not "demonstrate a relationship between tPCB exposure and reduced reproductive output". The assignment, then, is not objective: lack of a finding is not cause for a reduced ranking. GE's assertion that robins do not appear to be affected, within the test ranges measured, is supported by the ARCADIS study. The ERA secondarily sites the "nature of the study design … is not suited to assessing stressor-response relationships". That there are differences in the strengths of the two study in terms of control over where nest sites were located (we are not provided a figure on where the monitored nests are in the ARCADIS study – something that should be corrected in the next draft), the number of analyses on eggs and fledglings, the relative statistical power associated with the increased number of samples, is clear, and may provide better strength to the ranking. However, the ERA has already assigned the GE-study a "high value" for the Degree of Association.

RESPONSE 3.4-TT-11

The principal reason for the moderate weight assigned for the stressor/response attribute for the American robin field study was the small sample size available to estimate exposure to eggs and nestlings. This small sample size limited the dose-response evaluation, due to lower statistical power and reduced representation of exposure levels, in the tests of whether stressor/response relationships existed between exposure and the effects metrics included in the study. As a result, the only type of analysis that could be conducted for each of the effects metrics included in the study was to test whether significant differences existed between reference and target locations. This rationale will be made clearer in the revised ERA.

It is my opinion that the modeled exposure and effects endpoint was inappropriately assigned a moderate weighting value. The strength of any model lies in the correlation of predicted and observed effects. Table G.4-7 and G.4-8 show that the modeled exposure predicted a high magnitude of risk for tree swallows, and an intermediate risk for robins. Both field studies demonstrated, and were appropriately assigned low magnitudes of risk. The conclusion is that the model over-predicts risk for these species, and those should be assigned a low weighting value. The moderate assignment may be a residual of an a priori assignment given within the Work Plan, which stated:

> "Assuming that reasonable input parameters and distributions are applied to the modeling effort, a Medium weight-of-evidence will be given to this measurement endpoint." (Weston 2000, Section 7, page 7-76, lines 7 – 8)

RESPONSE 3.4-TT-12

Lack of concurrence between the results of the field studies and the modeling line of evidence should not be presumed to indicate that the field studies are correct and that the modeling results are incorrect. As discussed in the response to General Issue 8, each of the available lines of evidence has its advantages and disadvantages in terms of understanding and characterizing risk. It is for this reason that a WOE assessment is conducted. Thus, lack of concordance between lines of evidence is not a justification for down-weighting any particular line of evidence. Rather, lack of concordance is considered by EPA to be indicative of uncertainty regarding magnitude of risk posed by a stressor to the receptor of interest.

The weights assigned to the modeling line of evidence were not based on an a priori assignment in the Work Plan. Weights often need to be adjusted after studies or analyses are carried out because of changes in study design (e.g., due to unexpected weather events, lessons learned from earlier studies), difficulties in study execution, availability of improved tools or knowledge. Thus, the weights assigned in the risk characterization were done using the attributes for the specific endpoint.

3.4(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?

Valery Forbes:

Extrapolation to other species seems to have been performed as a qualitative comparison that only incorporated differences in exposure. It would be helpful if some argumentation could be provided to indicate whether these other species are likely to be more/less/ or equally sensitive to PCBs. Because low risk to insectivorous birds was indicated in the PSA, no risk assessment was performed for birds outside of the PSA.

RESPONSE 3.4-VF-7

As stated in Section G.4.6 of the ERA, "Effects data are not available for other insectivorous bird species living in the Housatonic River area. As a result, the same effects data used to estimate effects to tree swallows and American robins would be used for other insectivorous bird species." Other comparable insectivorous birds common to the area include the bank swallow, northern rough-winged swallow, barn swallow, cliff swallow, chimney swift, common nighthawk, eastern kingbird, and eastern phoebe. A qualitative assessment of risk was carried out for each of these species and a conclusion of species sensitivity in relation to tree swallows was provided in Section G.4.6.1.1.

Thomas W. La Point:

Given the limitations noted above, yes, they were objectively applied.

James T. Oris:

Yes.

Mary Ann Ottinger:

Birds were not considered at risk, based on tree swallow data. This conclusion follows the data collected, but may be an underestimate based on the choice of representative species and also does not consider pesticides or compounds that may be associated with industries, such as the paper mill. The survey of species in the downstream region should be revisited to verify this conclusion.

RESPONSE 3.4-MO-9

The tree swallow was selected as a representative species 6 years ago, before much of the current data were available that suggest that this species is not particularly sensitive to PCBs. Therefore, EPA agrees that risk estimates for the tree swallow may underestimate risks for other more sensitive insectivorous bird species. Please refer to the response to General Issue 14.A and Response 3.4-TT-5. Also, the purpose of the ERA is to determine risks due to releases from the GE facility, not from all sources. Thus, it is beyond the scope of this ERA to address risks due to other industries such as paper mills. In addition, EPA evaluated other COPCs (e.g., pesticides) in the PSA as part of the Pre-ERA and these were determined to not pose risk to wildlife.

Brad Sample:

A downstream evaluation not done for insectivorous birds. Rather extrapolation to other insectivorous bird species was provided instead. Some explanation should be provided to explain why downstream risks were not evaluated.

RESPONSE 3.4-BS-24

A downstream assessment was not conducted for insectivorous birds because risks in the PSA were low, and PCB and TEQ levels in soil and sediment are much lower downstream of Woods Pond than they are in the PSA. This explanation of why a downstream assessment for insectivorous birds was not conducted will be provided in the revised ERA.

Ralph G. Stahl:

The risk estimates were appropriately derived for reaches of the river where site-specific studies were not conducted.

Timothy Thompson:

Derivation of the upper and lower effects thresholds for both Aroclors and TCDD-TEQs (Figures G.3-1 and G.3-2), while conservative, are consistent with both scientific and ERA practice. Application of those thresholds to the lower reaches of the Housatonic, based on these conservative thresholds, show low risk and are appropriate.

Again, would recommend applying the MATC "rules" established for benthic infauna here as well. In the case of insectivorous birds, however, I do not believe that will significantly alter the results.

RESPONSE 3.4-TT-13

Please refer to the response to General Issue 6.

3.4(j) In the Panel members' opinions, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?

Valery Forbes:

The ERA concluded that risk to insectivorous birds is low, but that confidence in this conclusion is not high (does that mean low or moderate?). I concur that the risks to this receptor are low, and I would put moderate confidence on this conclusion, largely on the basis of the field studies.

RESPONSE 3.4-VF-8

EPA will reevaluate this wording in the revised ERA.

Thomas W. La Point:

Please refer to my response in 4.h. No, I do not think the conclusions of low risk to tree swallows are supported by the results. I would have concluded "moderate-to-high" risk, after reviewing the data. I have not put much weight (either way) on the robin study. In my opinion, that study should form the basis for a second study on robins, conducted over a minimum of two or three years.

RESPONSE 3.4-TL-10

EPA believes that adequate information exists to make an informed management decision regarding risk and that additional long-term studies, while they would be informative from a toxicological perspective, are not necessary to achieve this objective. There is a difference of opinion among reviewers regarding the WOE for the field studies for this and other endpoints. The WOE analysis will be revisited in the revised ERA. Please see response to General Issue 8.

James T. Oris:

Yes, supported.

Mary Ann Ottinger:

No. Data from other species would be required to make a valid conclusion of low risk to passerines and ground dwelling insectivorous species. In addition, there were data from wood ducks and mallard ducks collected from Reaches 5 and 6 that had moderate tPCBs (7.7-6.1ppm and 4.99-15.3ppm in the breast tissue, respectively; section 6.4.7) and low levels of pesticides. Opportunistic collection of house wren eggs (n=5; Reach 5) contained significant tPCBs levels and pesticide residues. Similarly, chickadee eggs (n=3) had tPCBs at moderate levels and pesticides.

RESPONSE 3.4-MO-10

In response to Reviewer comments, EPA is adding a duck species (wood duck) to the revised ERA, and has recently completed collection of a limited number of wood duck eggs from the PSA for analysis of contaminant concentrations. Given the lack of effects data for house wrens and chickadees, EPA is not considering adding these species to the revised ERA.

Brad Sample:

Although the information in the ERA does support the final conclusions, some revision and clarification is needed to indicate that the evaluation is for the local population and not for individuals.

RESPONSE 3.4-BS-25

The conclusions regarding risks of tPCBs and TEQ to insectivorous birds are for individuals in the locally exposed sub-population. Please refer also to the response to General Issue 9.

Ralph G. Stahl:

The information provided in the ERA supports the conclusion of low risk to local populations of insectivorous birds. Despite the statement that the conclusion is uncertain due to the lack of concordance between the lines of evidence, the 3-yr tree swallow study appears to be significantly robust and clearly supportive of the conclusion. The modeled exposure and effects results are interesting, but should be given a lower weight than was done in the weight-of-evidence analysis. If there is any uncertainty with this overall conclusion, it appears to be quite low.

RESPONSE 3.4-RS-4

There is a difference of opinion among reviewers regarding the WOE for this and other endpoints. The WOE analysis will be revisited in the revised ERA. Please see response to General Issue 8.

Timothy Thompson:

For insectivorous birds, the lines of evidence and data presented support the conclusion of the ERA.

3.5 Piscivorous Birds

3.5.(a) Were the EPA studies and analyses performed (e.g., field studies, site-specific toxicity studies, comparison of exposure and effects) appropriate under the evaluation criteria, and based on accepted scientific practices?

Valery Forbes:

The studies basically consisted of modelling, and yes these were appropriate.

Thomas W. La Point:

Given the limited amount of data for piscivorous birds, the use of kingfishers and osprey was valuable and appropriate. The kingfisher field study, albeit conducted with the best of intentions, has very low sample size and hence a low "power of the test." As with the insectivorous birds (in Section 4, above), the life span of kingfishers and osprey are very long compared to the duration of the field study. Hence, effects would not be expected to have been measured, unless they were of an extreme, acute, nature. However, the studies were certainly conducted based on accepted scientific practices.

The EPA modeling studies, using literature data and threshold values, are also based on accepted scientific practices.

James T. Oris:

During the GE comments, the use of osprey as an endpoint was criticized. However, based on comments from other public speakers and discussion of the panel, I believe the use of osprey was appropriate. Other endpoints, studies, and analyses were appropriate.

Mary Ann Ottinger:

Studies on piscivorous/carnivorous birds included a belted kingfisher study, which is a productivity study with visual monitoring of the nests. There was a limited n (n=9; 3 nests depredated) and the study was conducted in one breeding season. There was no assessment of growth; reproduction did not differ with tPCBs content. The rationale for the osprey study was that "belted kingfisher and osprey are the only piscivorous birds common to the area". However, the osprey study is a model, based on the foraging area being the entire PSA. As such, the number of birds actually impacted is not clear. There may be other avian species appropriate for the ERA, which inhabit the PSA in larger numbers. Candidate species might include owls, hawks, gulls, herons, or perhaps a species of egret.

RESPONSE 3.5-MO-1

EPA conducted an extensive ecological characterization of the PSA (ERA Appendix A) and, based on an inventory of both resident and migratory bird species, concluded that belted kingfisher and osprey are the best representative species for the Piscivorous Birds endpoint.

Hooded mergansers, a fish-eating duck, occasionally nest in the PSA, but not in sufficient numbers to permit an in situ study and therefore provide no advantage over osprey. Common mergansers migrate through the PSA, and also use the open water areas of the river and Woods Pond in the winter. However, they were not observed nesting in the PSA.

Because there is a rookery approximately 5 miles from the PSA, the great blue heron was also considered as a representative species for this endpoint. EPA does not consider great blue heron to be an appropriate representative species for piscivorous birds because:

- There are likely only 3 to 5 herons that forage in the PSA above Woods Pond Dam.
- No reproductive data are available for these particular herons because it is impossible to know which nests in the nearby rookery belong to the herons that forage in the PSA.
- Reproductive data are only available for the entire nearby rookery which consists of approximately 28 nests -- greater than 80% of these birds do not forage in the PSA above Woods Pond Dam because herons are territorial feeders.
- Impacts on individuals that forage in PSA above Woods Pond Dam would be obscured by the greater than 80% of birds that are not being exposed to high concentrations of PCBs.

Another wading bird species, the American bittern, was included in the risk assessment for threatened and endangered species (Section 11, Appendix K). Hawks and owls, while commonly observed in the PSA, are not considered to be piscivorous bird species and thus were not considered representative species for this endpoint.

Brad Sample:

EPA analyses for piscivorous birds were appropriate and based on accepted practices. Both the kingfisher and the osprey are suitable receptors for the PSA.

Ralph G. Stahl:

There were no field studies, or site-specific toxicity studies conducted by EPA for this assessment endpoint. The exposure and effects were all based on modeling. The fact that all the EPA-sponsored work was based on modeling has injected high uncertainty into the analysis.

RESPONSE 3.5-RS-1

EPA used the results of the GE study of kingfisher in the weight-of-evidence assessment for this species. EPA believes that it would be difficult to conduct other field studies on belted kingfishers in the PSA that would add substantial new knowledge to that already obtained from the GE field study. Thus, EPA did not conduct additional field studies for belted kingfishers.

It was not possible to conduct a similar study for osprey, the other species selected to be representative of piscivorous birds in the ERA, because of their large foraging range and low population density in the Housatonic River area. The uncertainty associated with having only one line of evidence (i.e., modeling of exposure and effects) for this representative species was explicitly acknowledged in both Section 8 (Section 8.5.6) and Appendix H (Appendix H.4.6) of the ERA.

Timothy Thompson:

Yes – the modeled exposures were appropriate and are based on accepted scientific practices. Some reconciliation should be made on the conclusions of reproductive risks to osprey, with the field observation that sitings within PSA have increased in numbers from 10, to 190.

RESPONSE 3.5-TT-1

As indicated in the osprey species profile included in the Ecological Characterization (ERA Appendix A.1), those sightings are for the entire State of Massachusetts from 1970 to 1990, not for the Primary Study Area of the Housatonic River. Thus, the sightings are not inconsistent with the assessment of reproductive risk to osprey in the PSA.

3.5(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria, based on accepted scientific practices, and incorporated, appropriately in the ERA?

Valery Forbes:

The belted kingfisher field study was unfortunately conducted for only a single breeding season and suffered somewhat from small sample size. However as the only site-specific field data on piscivorous birds this information is an important contribution to the ERA.

RESPONSE 3.5-VF-1

EPA concurs with this comment. The belted kingfisher field study was given an overall weighting of moderate/high in the weight-of-evidence assessment for piscivorous birds in the current ERA (Table H.4-4). As with all of the field studies, EPA will reconsider the weighting assigned to this study in the revised ERA. Also, please refer to the response to General Issue 8.

Thomas W. La Point:

See my response under 5.a. Conducting a study of this magnitude takes a lot of effort, time and personnel. However, for a relatively long-live species, to determine reproductive or growth effects with a one-season study produced results of limited value. Monitoring the burrows for kingfishers is laudable! However, very little, if anything, can be concluded from the results of six non-depredated clutches. This field study should be given little weight in the overall ERA.

RESPONSE 3.5-TL-1

EPA will reconsider the weighting assigned to this study in the revised ERA. Also, please refer to the response to General Issue 8.

James T. Oris:

GE study was, again, very limited in scope and suffered from serious limitations. No reference areas were included, no clutch sizes or hatching success were noted, the range of PCB concentrations was too narrow to explore a dose-response, and the sample size was too small to make any conclusions about the study. This was an admirable attempt at a difficult study, but it does not add a lot of meaningful information to the risk assessment.

RESPONSE 3.5-JO-1

EPA will reconsider the weighting assigned to this study in the revised ERA. Also, please refer to the response to General Issue 8.

Mary Ann Ottinger:

Although the Belted Kingfisher study was well conceived, with excellent observation techniques, there were a very small number of nests and no reference nests. Exposure was estimated by analysis of fish from the areas of the nest, which is an indicator of fish fed to the young. This results in making some aspects of the kingfisher study difficult to interpret. Steps to take include: clarify the number of nests in each area and relate to tPCBs of fish in the area of fishing by the parents, determine if there are any exposures that could qualify as reference area nests, and possibly include the additional nest observed by EPA contract staff (if data have been collected in a similar manner). Alternatively, a study with another avian species would address these issues.

RESPONSE 3.5-MO-2

EPA will reconsider the weighting assigned to this study in the revised ERA. Also, please refer to the response to General Issue 8.

The location and number of kingfisher nests in the PSA was determined by the GE field study, and confirmed by EPA (EPA found one additional nest). The information was included in the GE field study report and is summarized in Appendix H.4.2.2 of the ERA. No reference nests were included in the GE field study, nor have reference nests been monitored by EPA. It is not possible to conduct additional avian field studies in support of the ERA because of practical limitations.

The evaluation of exposure concentrations in fish for individual nests cannot be tied to specific fish tissue samples, but only estimated on a reachwide basis.

Brad Sample:

GE should be commended for attempting to collect field data on the resident kingfisher population in the PSA. The methods used are suitable and appropriate for this species. Unfortunately the study suffers from small sample size and limited study duration. As a consequence, utility of conclusions are limited - definitive conclusions about the presence or absence of effects cannot be made. Limitations to the conclusions include: 1) Data were from only a single year, with the representativeness of conditions in the year being unknown. 2) No data from a reference location were available for comparison. 3) low densities resulted in sample sizes being very small. The only conclusion that can definitively be made is that kingfishers do reside in the PSA and they appear to be reproducing. Collection of eggs, nestlings, or stomach contents for residue analyses would have greatly strengthened the data from this study. In the absence of these data, the magnitude of exposure experienced by birds at each nest is uncertain.

Comparatively high predation rates observed may be suggestive of an effect on nest attentiveness (may also be noise). Note that habitat in area is not of particularly high quality - maximum HSI was only 0.5. Note also that in comparison to nest success (after removal of depredated nests) from other areas, nestling survival from Housatonic is one bird fewer per nest - not statistically significant, but considering the small sample size and limited time duration, this result does raise questions.

In the WoE EPA considers the results of this study to be indicative of no evidence of harm. Given the limitations of the study, I believe it is more appropriately considered to be inconclusive (i.e., undetermined) in terms of evidence of harm.

RESPONSE 3.5-BS-1

EPA will reconsider the weighting assigned to this study in the revised ERA. Also, please refer to the response to General Issue 8.

Ralph G. Stahl:

The belted kingfisher study conducted by GE was appropriate under the evaluation criteria and represents the only field study conducted on this assessment endpoint. Unfortunately this study was conducted for only 1 year, the number of nests examined small, and no nests remaining in the examination were located in a reference area. Despite the small number of burrows (nests) that were evaluated for this study, the study was incorporated appropriately into the ERA. The belted kingfisher study was based on accepted scientific practices but could have been improved significantly by extending the duration and spatial scale. The loss of nests to depredation and other factors is unfortunate but does not warrant the exclusion of this study from the ERA.

Timothy Thompson:

I agree with my colleagues in noting that the overall study design and techniques were very good, and remarkable in their presentation. As noted above, the study at best demonstrates that reproducing kingfishers exist within the PSA, and contributes little beyond that. The few number of sampled nests, and the single season, contribute to very low power associated with the conclusions. Having said that, the kingfisher study is the only site-specific field data for this assessment endpoint, and is appropriately incorporated into the ERA.

3.5(c) Were the estimates of exposure appropriate under the evaluation criteria, and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?

Valery Forbes:

Yes seems so.

Thomas W. La Point:

Yes. I think the use of fish –rather than the limited data available for crayfish – was appropriate. In other situations where crayfish are more common, their use as prey items would be more important in estimating dietary intake. Given the nature of the Housatonic, and the fish community, the "substitution" of fish for crayfish is acceptable.

RESPONSE 3.5-TL-2

As described on pages H-13 to H-14 of the ERA, crayfish were included in the exposure model for belted kingfishers in Reach 5. The proportion of crayfish in the diet of kingfishers in Reach 5 had a mean of 14% and a range of 3 to 30%, which approximately corresponds to what has been observed in kingfisher dietary studies conducted in other riverine systems.

For the Woods Pond and Threemile Pond exposure models, crayfish were not included in the kingfisher diet because:

- Aquatic vegetation at these locations would conceal crayfish from kingfishers.
- Cyprinids and centrarchids of appropriate forage size are abundant at these locations and would be more visually attractive to foraging kingfishers.

James T. Oris:

See comments in 3.4. Also, the EPA should consider using exposure estimates as a distribution of concentrations based on the size distribution of prey items, instead of a single point estimate of exposure.

RESPONSE 3.5-JO-2

EPA believes this suggestion is worth exploring and will include the following in the revised ERA:

- A review of the literature to determine the preferences for certain fish size classes over others for piscivorous birds and mammals.

- The results of a sensitivity analysis to determine how this change in exposure concentrations would affect the risk estimate.

Mary Ann Ottinger:

Belted Kingfishers were studied to determine productivity, nest density; COCs were determined from the fish in the nest and estimates from samples collected in the area of the nest. The osprey model was generated from estimates of fish in the PSA and therefore may not be verified by data collected within the ERA. Because these species are more wide ranging in their foraging, it is possible that they routinely encounter the COCs in a patchy manner. Likewise, these birds are also likely to contact other contaminants, including pesticides applied to the agricultural fields, weed killers, and other environmental chemicals. Therefore, it is appropriate to consider other COCs for these species.

RESPONSE 3.5-MO-3

Other COPCs were screened out either in the Pre-ERA phase (see ERA Appendix B) or because their measured concentrations in the PSA were overestimated due to documented laboratory interference. Therefore, exposure to other COPCs in the PSA is considered minimal and should not pose a risk to wildlife. EPA agrees that osprey could encounter other COPCs outside of the PSA because they forage widely. However, the scope of this assessment is limited to the incremental risk due to contaminants in the Housatonic River as a result of direct and indirect (e.g., leaching from landfills) releases from the GE facility in Pittsfield, MA.

Brad Sample:

This comment applies to Appendices H, I, and K and relates to fish size in exposure analyses for piscivores. Piscivorous wildlife generally select fish opportunistically based on size. Exposure for kingfisher, osprey, bittern, mink and otter were all based on the UCL95 concentrations in fish that fell within the appropriate size classes. Although the largest size fish in the size range consumed by each species are likely to have the highest PCB concentrations, it is unlikely that the diet of a piscivore will consist extensively of the largest fish they can consume. Rather, they will consume fish that tend toward the center of their respective size range (these will have lower PCB concentrations). Using the UCL95 for all fish within the size range used by a species may overestimate exposure. [Another issue is that we do not know what the size distribution for fish within each range was - was the distribution even over sizes or biased? This will also affect the accuracy of the exposure estimate] An alternative would be to segregate fish into size classes, calculate UCL95s by size class, then incorporate the appropriate size class concentrations into the exposure estimate for each species.

RESPONSE 3.5-BS-2

EPA believes this suggestion is worth exploring and will include the following in the revised ERA:

- A review of the literature to determine the preferences for certain fish size classes over others for piscivorous birds and mammals.

- The results of a sensitivity analysis to determine how this change in exposure concentrations would affect the risk estimate.

Habitat used by osprey. It appears that this species was assumed to be exposed to PCB concentrations in fish from throughout the PSA. This does not seem appropriate. The upper sections of the PSA are generally too overgrown and shallow to allow significant use by the large piscivorous birds. The exposure estimation would be more appropriate if it weighted the exposure to be higher in the southern areas in the vicinity of Woods Pond and lower for the upper areas.

RESPONSE 3.5-BS-3

EPA agrees with the comment that osprey would spend more time foraging in the southern portion of the PSA (Reaches 5B, 5C, and 6) than they would in Reach 5A, even though there is suitable foraging habitat in all reaches. Therefore, EPA will exclude Reach 5A from the risk assessment for osprey in the revised ERA.

References for sources of parameter values in Tables H.2-9 and H.2-10 are lacking and should be provided.

RESPONSE 3.5-BS-4

EPA agrees with this comment, and will add references to the appropriate tables in the revised ERA.

It is not clear from where the fish (and crayfish) used to estimate exposure to kingfisher and osprey were collected. Reach 5 is a very large area. To understand how representative the exposure data are, some reference to the spatial locations addressed should be provided. Maps displaying sampling locations (an associated PCB concentrations in sediment) for biota used in the piscivore exposure would help.

RESPONSE 3.5-BS-5

Maps showing sample locations for fish and crayfish used in the exposure analyses will be provided in the revised ERA. Please refer also to the response to General Issue 1.

Fish were collected throughout Reach 5; fish sampling locations are shown on Maps 3-1 to 3-4 in Attachment J to the Ecological Characterization of the study area (ERA Appendix A).

Ralph G. Stahl:

The estimates of exposure were based solely on a modeling evaluation. These estimates are therefore highly uncertain and potentially problematic with respect to conclusions on risk. In this case however, the estimates are reasonable but not fully objective.

RESPONSE 3.5-RS-2

The risk assessment for belted kingfishers was based on two lines of evidence (modeling of exposure and effects and GE field study). The assessment for osprey was, however, based solely on the results of modeling of exposure and effects. The uncertainty associated with having only one line of evidence (i.e., modeling of exposure and effects) in the osprey assessment was acknowledged in both Section 8 (i.e., Section 8.5.6) and Appendix H (i.e. Appendix H.4.6) of the ERA.

Timothy Thompson:

In general, yes. I would like to have seen more use made of the crayfish data within the modeling approach – particularly since crayfish are directly exposed to sediments. Use of the extensive fish data is appropriate.

RESPONSE 3.5-TT-2

As described on pages H-13 and H-14 of the ERA, crayfish were included in the exposure model for belted kingfishers in Reach 5. The proportion of crayfish in the diet of kingfishers in Reach 5 had a mean of 14% and a range of 3 to 30%, which approximately corresponds to what has been observed in kingfisher dietary studies conducted in other riverine systems.

For Woods Pond and Threemile Pond exposure models, crayfish were not included in the kingfisher diet because:

- Aquatic vegetation at these locations would conceal crayfish from kingfishers.
- Cyprinids and centrachids of appropriate forage size are abundant at these locations and would be visually attractive to foraging kingfishers.

3.5(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria?

Valery Forbes:

These were largely literature based and no controlled toxicity studies were performed for this ERA.

Thomas W. La Point:

The effects metrics identified were appropriate. Whether or not they were adequately assessed is questionable. The field study has limited use in this circumstance.

RESPONSE 3.5-TL-3

As part of the EPA response to comments of the Reviewers regarding the consistency of weight-of-evidence assessments, EPA will reconsider the weighting assigned to the kingfisher field study in the revised ERA. Also, please refer to the response to General Issue 8.

James T. Oris:

See comments in 3.4

Mary Ann Ottinger:

The metrics were appropriate, but limited as discussed above.

Brad Sample:

The effects data appear to be suitable and appropriate. EPA has conducted a very thorough review of the available toxicity data and has made a very strong effort to adequately represent the range of potential effects.

Ralph G. Stahl:

The effects metrics identified were appropriate under the evaluation criteria.

Timothy Thompson:

Yes.

3.5(e) Were the statistical techniques used clearly described, appropriate, and properly applied for the objectives of the analysis?

Valery Forbes:

Yes.

Thomas W. La Point:

Yes.

James T. Oris:

See comments in 3.4

Mary Ann Ottinger:

The models were based on appropriate literature and as such should be reasonable. They should also consider the distribution of size of the fish that they consume and potential tPCB contents. There are some issues as discussed above in the conclusions presented in the two types of models that should be addressed.

RESPONSE 3.5-MO-4

EPA believes this suggestion is worth exploring and will include the following in the revised ERA:

- A review of the literature to determine the preferences for certain fish size classes over others for piscivorous birds and mammals.
- The results of a sensitivity analysis to determine how this change in exposure concentrations would affect the risk estimate.

EPA is uncertain as to what "two types of models" the Reviewer is referring to. For the tree swallow assessment, EPA used both a total daily intake (TDI) modeling approach and a microexposure modeling approach to estimate exposure of tree swallows to tPCBs and TEQ. The TDI modeling was done to facilitate comparison of exposure estimates to other wildlife species, but was not used to estimate risk. The microexposure modeling results were used to estimate risks of tPCBs and TEQ to tree swallows. For belted kingfishers and osprey, however, the TDI modeling approach was the only approach used to estimate exposure.

Brad Sample:

Similar issue with transparency of statistical analyses as discussed for earlier sections, apply to this section too.

As in the insectivorous bird section (comment 4.e above), it is not immediately clear how chemical concentration data were included in the exposure models. However, at the December meeting, Dwayne Moore stated that chemical concentrations were included in the Monte Carlo analyses as a point value (the 95% upper confidence limit on the arithmetic mean [UCL95]). I am concerned that this approach does not adequately nor accurately represent the variation and uncertainty in exposure. The exposure distribution that is generated only represents the uncertainty and variability about the distributions for the life-history based parameters (body weight, diet composition, food ingestion, etc.). Variability and uncertainty associated with PCB/TEQ concentrations are not represented. Depending on how large the sample size is and the variance associated with the data, use of the UCL95 may over represent exposure. Although the probability bounds analyses captures what should be all possible exposures, it is a very broad brush. The Monte Carlo analyses should provide what is believed to be the reasonable, most-likely exposure distribution based on expected distributions of all parameters.

RESPONSE 3.5-BS-6

Please refer to the response to General Issue 10.

Statistical analyses should be provided for data presented in Tables H.2-4 and H.2-5 plus Tables H.2-12 and H.2-13 to indicate which sites differed from each other.

RESPONSE 3.5-BS-7

EPA will include the results of the suggested analyses in the revised ERA.

Ralph G. Stahl:

The statistical techniques were clearly described and appropriate for the objectives of the analysis.

Timothy Thompson:

Within the limits of previous comments made about needing better descriptions of statistical methods, yes.

3.5(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria?

Valery Forbes:

Yes, though it has to be recognized that the risk characterization is based on more limited information compared to some of the other target species. Basing risk estimates on models rather than site-specific empirical information is, and should be, more conservative and tend to overestimate the risks. However it needs to be articulated that indications of high risk from models should not be interpreted in the same way as indications of high risk from sitespecific field studies. In my view the latter should trigger management actions, whereas the former should trigger collection of site-specific information to refine the risk assessment.

RESPONSE 3.5-VF-2

There is no reason to assume, in the absence of additional evidence, that the results from the field studies (particularly with the limitations identified) are "correct" and that lack of concurrence of these results with the modeling line of evidence indicates that the latter must be incorrect. As discussed in the response to General Issue 8, each of the available lines of evidence has its advantages and disadvantages in terms of understanding and characterizing risk. It is for this reason that a WOE assessment is conducted. Thus, lack of concordance between lines of evidence is not a justification for down-weighting any particular line of evidence. Rather, lack of concordance is considered by EPA to be indicative of uncertainty regarding the magnitude of risk posed by a stressor to the receptor of interest.

Thomas W. La Point:

Yes, the characterization of risk was appropriate for the kingfisher, based on the modeled exposure and effects. The variance and "clouding" of results provided by the admirably attempted field study does not provide information useful in determining effects on the kingfisher population. The field study would either have to have been conducted over a longer time period, with marked individuals, or have been conducted during the one season with more than six clutches. I recognize the difficulty in conducting such field experiments and also recognize the ease of criticizing them; I am not in any way minimizing the effort it took to gather the data. Nor am I denigrating the work conducted: it was conducted with the best of intentions. However, because of the limited number of burrows, and then some of those burrows depredated, not much can or should be made of the results.

All this supports my conclusion that the WOE risk categorization of "high" evidence of harm should hold for both species and overall for Piscivorous birds.

RESPONSE 3.5-TL-4

EPA will reconsider the weighting assigned to this study in the revised ERA. Also, please refer to the response to General Issue 8.

James T. Oris:

Yes, appropriate.

Mary Ann Ottinger:

The risk characterization is in line with the models, except that the osprey models are based on the literature and on the measured tPCBs in the fish that constitute their prey.

Brad Sample:

The risk characterization appears to be generally.

Ralph G. Stahl:

The characterization of risk is dependent on the single, limited duration, limited scope field study in the belted kingfisher. The remainder of the information used to characterize risk stems directly from modeling efforts. Although there is great uncertainty in the conclusions drawn on risk, the characterization of risk is supported by the available information and was appropriate under the evaluation criteria.

The modeled exposure and effects, given the weaknesses found in the belted kingfisher study, were given an appropriate weight in the characterization of risk.

Timothy Thompson:

Within the context of the limited information presented in the ERA (modeling and low power with the kingfisher study), the assessment of risk appears to be appropriate.

3.5(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made.

Valery Forbes:

The uncertainties would be reduced by having additional field-based information on exposure or effects for this receptor, particularly given the large discrepancy between the modelled magnitude of risk (high) and that based on the field study (low).

RESPONSE 3.5-VF-3

EPA agrees that more field-based information would be useful to resolve the discrepancy in results for the two lines of evidence used in the belted kingfisher assessment. Additional field studies in support of the ERA are not possible due to practical limitations. The uncertainty that arises because of the discrepancy between the two lines of evidence is acknowledged in the ERA in both Section 8 (Section 8.5.6) and Appendix H (Appendix H.4.6).

Thomas W. La Point:

Yes, the uncertainties were acknowledged. The only improvement I can suggest is to include other –cageable- piscivorous birds (such as mergansers) and, if another field study were to be performed, to recommend that the study duration be as long as possible, well over two years and using banded birds. The use of a surrogate species, like a fish-eating duck, to directly study PCB uptake would lessen the need for prey PCB burden estimates. Also, should the caged ducks be studied over one reproductive season, the maternal burden provided to eggs or nestlings could be accounted for.

RESPONSE 3.5-TL-5

EPA agrees that additional field studies would reduce uncertainty and improve understanding of risk for piscivorous birds. A limited field effort was performed in May 2004 in response to Reviewer's comments to collect wood duck eggs and analyze them. The results will be discussed in the revised ERA.

James T. Oris:

Uncertainties were appropriately described.

Mary Ann Ottinger:

There were uncertainties in the data for the belted kingfisher due to small n of nests and there were no data collected in the PSA for osprey, resulting in a model-based approach for the ospreys. The inclusion of dose response data from laboratory studies is valuable and there is no reason to disregard data from any species, including the domestic chicken unless the sensitivity estimate is available for the species under study. Furthermore, exclusion of data, such as those from domestic poultry will potentially underestimate the risk for other species.

RESPONSE 3.5-MO-5

EPA agrees with this comment. Nearly all test species used in human and ecological risk assessment have gene pools that are relatively homogeneous in comparison to wild species. Chickens are no different from the usual test species in this respect. Also, only a limited number of avian species have been evaluated for effects from tPCBs; therefore, a threshold range approach is necessary to encompass the range for all wild species that could occur in PSA. In using this approach in the ERA, all appropriate test data were included in the derivation of a threshold range.

Brad Sample:

Although the uncertainties are generally well captured in the ERA report, the relative importance of various uncertainties is not discussed. Interpretation and understanding of the results would benefit by such an analysis. Spatial heterogeneity of contamination and the degree to which the exposure data actually represent the PSA are very significant issues. These need to be discussed.

RESPONSE 3.5-BS-8

EPA agrees with this comment, and a discussion of the relative importance of the sources of uncertainty will be provided in the revised ERA.

Ralph G. Stahl:

There are significant uncertainties in the analysis of the assessment endpoints, particularly that of the osprey. The information supplied by GE during the January 2004 public meetings suggest that osprey are not found in the Housatonic River watershed, and, therefore, should not be used as a potential receptor in the ERA. This point does not seem to be covered adequately in the ERA and should be examined in view of the potential significance of a piscivorous bird receptor to the final risk management decision.

RESPONSE 3.5-RS-3

EPA believes that osprey are an appropriate representative species for piscivorous birds because:

- Ospreys have expanded throughout appropriate coastal area habitats and are now moving inland.
- The PSA has appropriate habitat for ospreys.
- One pair of osprey displayed courtship behavior during the breeding season in the PSA during the EPA field investigation.
- There is a high likelihood that ospreys will be nesting in the PSA in the near future as they continue to recover from exposure to DDT and other organochlorine pesticides.

Timothy Thompson:

Uncertainties in the modeled projections were generally described. The low power associated with one season and 6 nests within the kingfisher study should be more clearly articulated in the final document.

RESPONSE 3.5-TT-3

EPA concurs with this comment and believes that this information has already been clearly articulated in the ERA (see page H-49 and Table H.4-3).

3.5(h) Was the weight of evidence analysis appropriate under the evaluation criteria? If not, how could it be improved?

Valery Forbes:

This was limited due to lack of data.

Thomas W. La Point:

Please refer to my response under 5.f., above. I do not think the WOE analysis is appropriate for these species, as the study duration was not long enough, if the field study is to be given weight in the ERA conclusions. The modeling study provides a more conservative estimate, yes, but this (and the conclusions stemming from accepting the higher degree of risk) should be weighed against the loss of the avian resource. In Table 8-5.3, in Volume 2, Section 8, the modeling results indicate risk is "high" for the kingfisher and osprey. The weighting of the results (Tables H 4-6 and 4-7) indicate "moderate weighting." The field study indicates "low" risk. However, the field study, unfortunately, is simply not useful in determining long-term effects.

RESPONSE 3.5-TL-6

EPA will reconsider the weighting assigned to this study in the revised ERA. Also, please refer to the response to General Issue 8.

James T. Oris:

See previous comments on WOE analysis. The weighting used for the kingfisher study were too high. EPA should reconsider weighting assigned to this study.

RESPONSE 3.5-JO-3

EPA will reconsider the weighting assigned to this study in the revised ERA. Also, please refer to the response to General Issue 8.

Mary Ann Ottinger:

The WOE shows potential risk, but the interpretation is somewhat inconclusive due to the limited measures conducted on the animals collected. It would have been valuable to have both historic data on the species in the PSA as well as studies that were conducted over several years.

RESPONSE 3.5-MO-6

EPA agrees that historical data would have been useful; however, there were no comprehensive studies performed on species in the PSA in previous years to use as a reference.

Brad Sample:

The weight-of-evidence (WoE) analyses is generally appropriate but is not adequately transparent.

RESPONSE 3.5-BS-9

Please refer to the response to General Issue 8.

The WoE analyses use best professional judgment to apply final weightings for the various lines of evidence and their attributes. It is unclear how final weightings were derived. Some discussion as to how these weights were derived needs to be provided – were all attributes judged equally?

If not need to know why and this needs to be explained – this will allow reviewers to judge the process.

RESPONSE 3.5-BS-10

Please refer to the response to General Issue 8.

Weighting for field study appears high – given the limitations of the study, I would rank it as moderate at best. In addition, as stated above, I believe it is more appropriately considered to be inconclusive (i.e., undetermined) in terms of evidence of harm.

RESPONSE 3.5-BS-11

EPA will reconsider the weighting assigned to this study in the revised ERA. Also, please refer to the response to General Issue 8.

Ralph G. Stahl:

The weight of evidence analysis was appropriate under the evaluation criteria. As noted earlier, the field study provided a limited dataset and was weighted appropriately in this analysis. Thus the modeled exposure and effects assessment needed to receive a weighting of "moderate".

Timothy Thompson:

For osprey, it may have been more appropriate to divide the PSA into north and south based upon the river characteristics. The narrow, relatively fast moving water is not going to be conducive to osprey hunting, which is more likely to occur within the ponded areas in the south. While this may seem inconsistent with the rest of the document, from an exposure standpoint it makes sense.

RESPONSE 3.5-TT-4

Please refer to Response 3.5-BS-3.

3.5(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?

Valery Forbes:

Not performed.

Thomas W. La Point:

As there are no data for these risk estimates, the modeling results should hold (as "high" evidence of harm, with moderate weighting). As uncertainty in the data provide for a conclusion of "low" risk, I disagree with this conclusion. Uncertainty, with these two species and given the limited information on piscivorous birds, should lead to more conservatism.

RESPONSE 3.5-TL-7

EPA will reconsider the weighting assigned to this endpoint in the revised ERA. Also, please refer to the response to General Issue 8.

James T. Oris:

Yes. However, the uncertainty of using literature derived MATC values should be discussed. Conclusions based on these estimates are not site specific and have high uncertainty.

RESPONSE 3.5-JO-4

EPA did not derive an MATC for piscivorous birds, as was done for mink, river otter, and bald eagle. If, however, the Reviewer is referring to the threshold ranges that were derived for tPCBs and TEQ in the effects assessment, then EPA concurs that such ranges are uncertain because they are not species- or site-specific. This source of uncertainty was acknowledged in the "Sources of Uncertainty" sections in Section 8.5.4 and Appendix H.4.4 of the ERA.

Mary Ann Ottinger:

Eagles identified as potentially at risk, based on the information on sediment tPCBs and on the potential sensitivity of these species to contaminants.

Brad Sample:

A downstream evaluation not done for piscivorous birds. Rather extrapolation to other piscivorous bird species was provided in Appendix K instead. Some explanation should be provided to explain why downstream risks were not evaluated.

RESPONSE 3.5-BS-12

A downstream assessment was not conducted for belted kingfisher because the weight-of-evidence conclusion was that this species was at low risk in the PSA. Risks to kingfisher are likely to be much reduced further downstream because tPCB and TEQ concentrations in fish are markedly lower below Woods Pond. Thus, a downstream assessment was judged to be unnecessary.

A downstream assessment was not conducted for osprey because the conclusion of high risk in the PSA was based on one line of evidence and had higher uncertainty. Further, the effects metric was not species-specific, adding further uncertainty to the assessment conclusion. EPA will reevaluate the appropriateness of a downstream estimate of risk for the osprey and provide further discussion for both species in the revised ERA.

Ralph G. Stahl:

The risk estimates were objectively and appropriately derived for reaches of the river where site-specific studies were not conducted.

Timothy Thompson:

Within the limits of uncertainty associated with model projections, yes.

3.5(j) In the Panel members' opinions, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?

Valery Forbes:

Yes.

Thomas W. La Point:

No, for the reasons detailed in Sections 5.h, I, and j, above.

RESPONSE 3.5-TL-8

Please refer to Responses 3.5-TL-6 and 3.5-TL-7.

James T. Oris:

Within the limits of the available information used in this section, the conclusions are generally supported. Some discussion is needed on the discrepancy between the modeled effects (high risk) and the field studies (low/no risk).

RESPONSE 3.5-JO-5

EPA provided this discussion in the ERA (Section H.4.6).

Mary Ann Ottinger:

There is a great deal of uncertainty due to the field study and to the limited information used to model the risk. The assessment remains somewhat inconclusive although it is reasonable based on the tPCBs in the prey. Ultimately, the real extent and potential for long-term impact is not clear for these species.

RESPONSE 3.5-MO-7

EPA concurs with this comment for both belted kingfisher and osprey. The sources of uncertainty indicated by the Reviewer have been acknowledged in Section 8 (Section 8.5.6) and Appendix H (Appendix H.4.6) of the ERA.

Brad Sample:

Although the information in the ERA does support the final conclusions, some revision and clarification is needed to indicate that the evaluation is for the local population and not for individuals.

RESPONSE 3.5-BS-13

The conclusions regarding risks of tPCBs and TEQ to piscivorous birds are for individuals in the locally exposed sub-population. Please refer also to the response to General Issue 9.

Ralph G. Stahl:

Based upon the information provided in the ERA, the evaluation supports the conclusion of low risk to local populations of belted kingfishers, and possibly other piscivorous birds. The conclusion of potentially high or moderate risk to osprey is supported by the modeling work, but is highly uncertain given the lack of field studies or biological survey data from the PSA on this species. In general the conclusions on risk drawn from the information in the ERA carry a high degree of uncertainty.

Timothy Thompson:

Within the limits of the information, these are the only conclusions one could draw.

3.6 Piscivorous Mammals

3.6.(a) Were the EPA studies and analyses performed (e.g., field studies, site-specific toxicity studies, comparison of exposure and effects) appropriate under the evaluation criteria, and based on accepted scientific practices?

Valery Forbes:

The mink feeding study was appropriately conducted. Although the field surveys seemed to find very few animals, this seems to be a common feature of these types of studies.

RESPONSE 3.6-VF-1

Although mink and river otter are not often observed directly in the field, they do leave tracks, scats, and other evidence when they are present in an area. The methods used in the field surveys conducted on behalf of EPA (e.g., snow-tracking and scent post surveys) have been shown to reliably indicate mink and river otter abundance at other sites. Thus, EPA believes that very few animals were found (particularly in spring and summer) because there were few animals to be found, not because of inherent limitations of the field survey methods. The field surveys conducted on behalf of GE also found little evidence that mink and river otter are present in the PSA during spring and summer. The latter surveys, however, did find evidence that mink and river otter are in the PSA during winter, as would be expected given that these animals expand their foraging ranges and/or seek new territories during winter. Little or no evidence was found in either the EPA or GE surveys that mink or river otter are resident in the PSA during the non-winter months. Additional discussion of this topic was presented in ERA Appendix I.4.2.

Thomas W. La Point:

Yes. The mammals selected for study, mink and otter, are very difficult to monitor in the wild. Hence, the reliance on modeling parameters and the mink feeding study (MSU) were highly appropriate. Although useful, the wide variances associated with locating, tracking, and identifying mink or otter during winter has less use – but was appropriately conducted.

RESPONSE 3.6-TL-1

Please refer to Response 3.6-VF-1 above.

James T. Oris:

Studies and analysis were appropriate. However, some of the data collected were not used and should be applied to the risk characterization. Specifically, jaw deformities and kit deformities should be accounted for in the risk assessment. These are significant impacts to the organisms and are predictive of long-term effects in the population. Again, I have the same comments concerning the definition of the assessment endpoint as for fish and birds.

RESPONSE 3.6-JO-1

Please refer to the response to General Issue 15.B regarding use of the data generated from the mink feeding study in the revised ERA. Please refer to the response to General Issue 2 regarding the definition of the Assessment Endpoint for piscivorous mammals.

Mary Ann Ottinger:

The representative species chosen were mink and otter. The MSU feeding study with mink fed fish from the PSA showed effects on the survival of kits in the 3.7mg/kg tPCB group at 6 weeks of age; jaw lesions were observed in kits in a dose-dependent manner. These additional measures are important for the ERA; specifically as there was a dose-dependent relationship observed in EROD and in jaw lesions in the kits. The significance of the lesions should be considered, i.e., are lesions associated with impaired immune response or localized tumorgenesis. The survey provided valuable information about the presence of mink and otter; scat samples gave information about prey.

RESPONSE 3.6-MO-1

Please refer to the response to General Issue 15.B.

Brad Sample:

The field and laboratory studies conducted by EPA were generally appropriate and acceptable. There however are several issues with both studies that should be addressed.

Mink Bioassay

The PSA fish incorporated into diet for the test mink were represented by large goldfish and carp. I recognize the need to have adequate biomass to create sufficient test diet to sustain the mink for the study duration. However, due to their size, large goldfish and carp are not fish that mink in the PSA would be consuming to any great degree. In addition, the feeding habits of goldfish and carp may potentially result in a different contaminant mixture than other fish. To address this uncertainty, I would recommend including an evaluation of how the contaminant mixture in the fish used for the diet compares to that for the fish assumed to represent the diet of mink in the PSA. Questions to include are, 1) how does congener mixture vary by fish size (age), and 2) how does congener mixture vary by fish species.

RESPONSE 3.6-BS-1

EPA evaluated the congener mixture in fish species and size classes, including the mixture used for the mink diet. The results demonstrated that the congener composition of the carp and goldfish used to prepare the diets for the mink feeding study was generally similar to the congener composition in freshwater fish in the PSA that would form the diet of the mink in the wild. A discussion of this analysis is presented in Section I.3.2.6 of the ERA. Although there is a reference to additional detail in Appendix C.7, this particular analysis was inadvertently omitted from the larger analysis of congener composition across media that was included in that appendix. Please refer also to the response to General Issue 15.B.

The test diets were only analyzed for chlorinated compounds – what about PAHs (might not be detected) or metals?

RESPONSE 3.6-BS-2

The test diets were analyzed for organochlorine pesticides, tPCBs, and non-ortho and mono-ortho PCB, PCDD and PCDF congeners. As discussed in ERA Appendix B, the data on organochlorine pesticides were not considered in the ERA or in the mink feeding study because these COCs were screened out using conservative screening criteria. Bioaccumulative metals were analyzed in a number of adult largemouth bass from the PSA to resolve the question of potential risk from metals to fish-eating receptors. The data were used in the screening of COPCs, and none of the metals exceeded the screening criteria. The screening analysis for bioaccumulative metals in fish indicated that concentrations of bioaccumulative metals were not of concern for mink. PAHs were not analyzed in fish.

DDE was elevated in the PSA fish used in the diet. This is a potential confounding factor that needs to be explained more. Potential effects that this may have had should be discussed. Concentrations of DDE in the final diets should be presented if measured. How do these concentrations relate to what is seen in other studies?

RESPONSE 3.6-BS-3

DDE and several other pesticides were retained as COPCs in fish tissue following the Tier III evaluation in the Pre-ERA. However, as was noted on page B-26 of the ERA, pesticide results may have been overestimated due to laboratory interference. EPA agrees that the data for other COCs should be better summarized and further discussed in the revised ERA.

Table 3 from Bursian et al. does not report PCB or TEQ concentrations – this would be helpful to see in relation the other organochlorines.

RESPONSE 3.6-BS-4

EPA agrees with this comment. The Bursian et al. report will be revised to include the data, and the data will also be provided in the revised ERA.

Just as a note, it is unfortunate that the dose range used in the study did not extend to one or possibly two higher doses. These higher doses would have likely produced more severe effects and would have strengthened the overall dose-response relationships.

RESPONSE 3.6-BS-5

EPA agrees that, in hindsight, knowing that the contaminant mixture present in the PSA appears to be less toxic than observed at other sites, higher treatment doses would have been useful. However, at the time of the study, the Principal Investigators at Michigan State University believed that a higher dose treatment would have led to mortality of adults, based on the results of previous mink feeding studies which they had conducted. Accordingly, they did not feel it was appropriate to include higher concentration treatments in the study.

Field Survey

Although my impression of the EPA study is that it is appropriate and well conducted, GE raised a valid issue at the January meeting: habitat differed between the PSA and the references areas. Whereas the PSA is a river and it's floodplain, the reference areas were all comprised of ponds and their shorelines. This issue should be directly addressed in the ERA. Potential habitat and methodological differences (transects in the PSA versus whole shoreline surveys in the reference areas), and how they may have affected the results should be discussed when these data are first presented and in the uncertainties section.

RESPONSE 3.6-BS-6

EPA does not believe that differences in mink and otter abundance between the PSA and reference areas are primarily the result of habitat differences. EPA will provide further discussion of this issue in the revised ERA.

Ralph G. Stahl:

The EPA field studies for mink were appropriate under the evaluation criteria and based on accepted scientific practices.

There is one area of potential concern regarding the mink feeding study and its acceptability under the evaluation criteria. It was mentioned at the January 2004 public meetings that kits which had died unexpectedly in the feeding study were not necropsied nor the cause of death determined. It is standard practice in long term studies, particularly where animals might die from natural causes and/or disease during long periods of confinement, to necropsy those animals which die unexpectedly and to the extent feasible, determine the cause of death. Without this approach significant uncertainty results in any conclusions regarding exposure to a toxicant and the survivability of the adults or offspring. Therefore, if EPA's contractor did not necropsy the animals which died unexpectedly, nor attempt to determine the cause of death, then the mink feeding study does not meet the evaluation criteria and is not based on accepted scientific practices.

RESPONSE 3.6-RS-1

All kits that died after 6 weeks of age were necropsied; the pathologist's assessment is contained in the mink feeding study report (all died of infections thought to be unrelated to PCBs). Kits that died prior to weaning at 6 weeks were not necropsied as per standard practice. Kits were handled only 3 times during this period (to obtain body weights at birth and at 3 and 6 weeks of age). Excessive disruption of the female can cause her to eat her young, particularly from birth to 3 weeks of age. If a kit is unhealthy or if it dies, the mother will often eat it, leaving no carcass to necropsy. If a kit dies and is not disposed of by the female, it usually is not a candidate for necropsy because of the length of time between death and discovery of the carcass, which often is buried in the bedding within the nest box. In addition to the mothers cannibalizing their young, the kits may cannibalize a weaker sibling as they approach 6 weeks of age.

While EPA appreciates the Reviewer's concerns, the necropsy of all kits that die prior to 6 weeks of age is simply not possible. EPA states in the report that body weights of kits in the high dose group at 3 weeks of age were significantly lower compared to controls, which implies a treatment-related effect on growth (a hallmark of PCB toxicity). This was followed by significant mortality from 3 to 6 weeks of age in the same dose group (as well as a return of average body weights in this treatment to control values). EPA and the Principal Investigator believe that those kits that were low in body weight at 3 weeks of age died during the subsequent 3-week period.

EPA does not agree with the Reviewer's comments that the study should be discarded because it was not possible to perform necropsies on kits less than 6 weeks of age. Previous mink studies have shown kit survivability to be a sensitive indicator of PCB toxicity in this species, and at least one study has demonstrated significant delivery of PCBs from female to kit during lactation (Bleavins et al. 1980; Heaton et al. 1995). Neither of these studies necropsied mink kits that died prior to 6 weeks of age. More importantly, deaths due to causes other than contaminants in the diet would be expected to randomly occur across dietary treatments. This was not the case because in the highest dose treatment, there was a significant effect on growth of mink kits from 0 to 3 weeks of age, and a significant effect on survival of mink kits from 3 to 6 weeks of age.

Timothy Thompson:

Yes, the combination of solid field studies, with the caged mink feeding exposure, is a solid piece of work that follows (frankly, exceeds!) general scientific/toxicological practice.

One issue to consider further is whether the model dataset was biased by only using carp and goldfish as the prey species. Mink are opportunistic, and will take a variety of prey species, but carp likely represent the highest levels of PCBs in fish in the River. A good discussion in the uncertainty section on this issue would probably address that concern.

RESPONSE 3.6-TT-1

Please refer to Response 3.6-BS-1 above.

Kudos for the congener work. That was excellent and helpful. Better use of the other organochlorine and metals data would also be recommended.

RESPONSE 3.6-TT-2

Metal concentrations were not measured in the dietary treatments of the mink feeding study (Bursian et al. 2003); however, bioaccumulative metals were analyzed in adult largemouth bass from the PSA to provide an estimate of the exposure of piscivorous receptors to these contaminants. No metals or organochlorine contaminants other than tPCBs or the TEQ congeners screened through as COCs for the wildlife assessments.

3.6(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria, based on accepted scientific practices, and incorporated, appropriately in the ERA?

Valery Forbes:

The field surveys would have been strengthened by inclusion of a reference site.

Thomas W. La Point:

Yes. The survey for mink and otter tracks (including scent posts, visible sightings, and foot track identifications) were appropriate. However, as with most initial surveys, the surveys provide good background for how to conduct further studies. Such studies, were they to be conducted, might include collaring selected adults and following them, incorporating GIS approaches to determine home range, migratory range, and den location(s). Similar studies (of radio-collared animals) have been conducted on mink and marten in the Wyoming and Colorado Rockies, with considerable success.

James T. Oris:

The GE studies in this section seem to have the most inherent limitations. Most of these were discussed in the ERA, but in question and answer sessions during the document presentation meeting and the public panel meeting, it became clear that the methodologies used by GE personnel were not all appropriate. These studies should be severely down-weighted or not used at all.

RESPONSE 3.6-JO-2

EPA will review the weights assigned to both the EPA and GE field studies for mink in the revised ERA. Please refer also to the response to General Issue 8.

Mary Ann Ottinger:

Field survey (2001-03) showed evidence of mink and otter in winter, but in low numbers at other seasons compared to informal data from sightings in the previous years. There is an uncertainty due to the sampling and observations.

Brad Sample:

GE conducted an independent field study to evaluate the presence and abundance of mink and river otter in the PSA. The EPA identified numerous limitations to the GE study (Appendix I, Pages I-62 and I-63). Based on my review of the GE report, I concur with the EPA comments. I think that the GE study greatly overstates the abundance of mink and otter in the PSA, and I think that the study was appropriately incorporated into the ERA.

Ralph G. Stahl:

The GE field study in mink was appropriate under the evaluation criteria, based on accepted scientific practices, and was incorporated appropriately into the ERA. There is one area of concern regarding this statement. Under questioning during the January 2004 public meetings it was revealed that standard practices for implementing scent posts for mink or other mammals may not have been followed consistently. For example, the ERA speculated that the lack of sightings at GE-implemented scent posts may have resulted from the posts being compromised by human contact (scent). To offset the potential for compromising the posts, field crews are generally instructed to wear rubber gloves and rubber boots. It appears that those who conducted the scent post studies for GE did not wear rubber gloves consistently during the studies. Whether this or another factor resulted in the "low" numbers of individual mink visiting the posts is unknown. If it is determined that there were numerous infractions with regards to the implementation of the scent posts by GE's contractors, then the mink field study would not meet the evaluation criteria.

RESPONSE 3.6-RS-2

The limitations regarding the GE field study noted in the ERA were based on observations made during EPA oversight of GE's mink field study. One of these observations was that study protocols were not always adhered to and may have compromised the survey results.

Timothy Thompson:

The same general comment raised previously concerning the GE studies are valid here. The work is solid and appears to have been lead by competent scientists. The study lacks a formal objective beyond demonstrating mink can be found within the PSA. This was clearly demonstrated, but not much else can be derived from their work. The issue of whether the GE field personnel were appropriate, and how gloves/rubber boots effect scent posts is well outside of my field of knowledge – I leave others to comment on this.

3.6(c) Were the estimates of exposure appropriate under the evaluation criteria, and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?

Valery Forbes:

As far as I can determine.

Thomas W. La Point:

The exposure period appears to account for seasonal variability in dietary composition for mink and perhaps less so for river otter. There is great value in using an exposure period that encompasses the reproductive cycles of these species. Although experimental, and controlled by the laboratory diet, results of dietary exposure were also well assessed in the mink feeding study. From Section 9, Volume 2 and Appendix I, it appears that the estimates of dietary composition (percentage of fish, invertebrates, etc) and daily intake were appropriate. One particularly good approach was to incorporate estimates of 10% and 100% foraging times in the PSA. These estimates provide a good "bound" for modeling exposures.

There is much more uncertainty in the estimates for the otter. However, without including a radio-collared study of otter migration, home range, and "local" behavior patterns, there will remain a large uncertainty (as expressed in this ERA) on otter exposure.

James T. Oris:

Some attention should be given to limitations of the feeding study. The MSU study used only goldfish and carp, but mink certainly eat more than these fish. This could have caused a bias in the PCB congeners in the exposure as well as other COCs. Some linkage with the analytical chemistry data is necessary here to alleviate concerns about the dosing regime.

RESPONSE 3.6-JO-3

EPA evaluated the congener mixture in fish species and size classes, including the mixture used for the mink diet. The results demonstrated that the congener composition of the carp and goldfish used to prepare the diets for the mink feeding study was generally similar to the congener composition in freshwater fish in the PSA that would form the diet of the mink in the wild. A discussion of this analysis is presented in Section I.3.2.6 of the ERA.

The test diets were analyzed for organochlorine pesticides, tPCBs, and non-ortho and mono-ortho PCB, PCDD and PCDF congeners. As discussed in ERA Appendix B and in Chapter 6, organochlorine pesticides were not considered further in the ERA because these COCs screened out using conservative screening criteria. Bioaccumulative metals were analyzed in a number of adult largemouth bass from the PSA to resolve the question of potential contribution to risk from metals to fish-eating receptors. The data were used in the screening of COPCs, and none of the metals exceeded the screening criteria. EPA agrees that the data for other COCs should be better summarized and further discussed in the revised ERA.

EPA needs to address GE's critique of no evidence of dose response in kit survival.

RESPONSE 3.6-JO-4

Please refer to the response to General Issue 15.B.

Mary Ann Ottinger:

The mink study was conducted in both the field and lab. Laboratory studies confirmed the impact of fish from the PSA on kits produced by exposed females. Field studies were less conclusive. Field data on otters (sightings) provided information on the presence of this species and likely prey in the areas of observation. Models developed should be reasonably accurate because the COCs would have been present in the laboratory study.

Brad Sample:

As discussed for avian piscivores in comment 3.5.c, the exposure analyses would be strengthened by increasing the resolution of the fish size-class data in the exposure model.

RESPONSE 3.6-BS-7

EPA believes this recommendation is worth exploring and will include the following in the revised ERA:

- A review of the literature to determine the preferences for certain fish size classes over others for piscivorous birds and mammals.
- The results of a sensitivity analysis to determine how this change in exposure concentrations would affect the risk estimate.

Ralph G. Stahl:

The estimates of exposure were appropriate under the evaluation criteria. As noted previously, it is important to provide and discuss all chemicals detected in fish tissues particularly since these substances become the basis for exposure evaluations in the mink feeding studies, and ultimately in the risk designation. The datasets for the fish used in the feeding studies appear to have been truncated and/or summarized so that only tPCBs and TEQ were reported. This is not appropriate under the evaluation criteria and should be rectified by including the specific information on fish tissues in the main text of the ERA.

RESPONSE 3.6-RS-3

The Pre-ERA described in Appendix B and the subsequent screening in Chapter 6 led to the screening out of all COPCs except tPCBs and TEQ for the wildlife assessments. However, EPA agrees that the data for other contaminants should be better referenced and discussed and will do so in the revised ERA.

Timothy Thompson:

My only comment here is that the final ERA account for GE's comment that there is no dose/response relationship in the reproductive caged-mink endpoints (see GE Figure 53). If there is no statistical differences between stations, and the relationship is not strictly dose dependent, than it is difficult to make a case that PCBs are impacting in the PSA.

RESPONSE 3.6-TT-3

Please refer to the response to General Issue 15.B.

3.6(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria?

Valery Forbes:

Yes, with the possible exception of the jaw lesions which I am not sure how to interpret in terms of impacts on mink populations. The suggestion that this effect could lead to starvation is speculation.

RESPONSE 3.6-VF-2

Please refer to the response to General Issue 15.B.

Thomas W. La Point:

Yes. The feeding study is particularly strong and results stemming from that study should (in my opinion) carry a lot of weight for the ERA results for piscivorous mammals. The studies cited in Appendix I indicate that mink are very sensitive to dietary PCB levels, confirming results of the feeding study in this ERA. There were, of course, no effects metrics quantified by the surveys; hence, these field studies were of much more limited use in determining effects on mink or otter. The fact that no mink were found (EPA study) in habitats appropriate for mink is somewhat disconcerting, and somewhat modified by the GE field survey. However, both studies are field surveys for species known to be extremely difficult to monitor in the wild, without radio collars attached. Hence, one would expect (at least this panelist would expect) few quantitative results on effects metrics to derive from field surveys.

The USEPA may gain some valuable insights into the NOEL for "aquatic mammals" in the recently-published (2001) NAS/NRC publication, "A Risk-Management Strategy for PCB-Contaminated Sediments." In it (pg.405), an NOEC for tissue residue total PCB was estimated to be 10 mg/g, lipid normalized. The number was cited from the work of Kannan, et alia, (2000).

RESPONSE 3.6-TL-2

EPA will review Kannan et al. 2000 and, if applicable, integrate information from this source into the revised ERA.

James T. Oris:

As mentioned in 3.6(a), there were effects measured, but not used. Several sublethal endpoints such as jaw lesions, could be used here. These provide an additional line of evidence, are indicative of delayed effects, and are pertinent to population health. Unfortunately, they are not addressed by the assessment endpoint as currently stated in the ERA.

RESPONSE 3.6-JO-5

Please refer to the response to General Issue 15.B.

Mary Ann Ottinger:

The laboratory studies with the mink included endpoints that were informative and were included as part of the assessment. Field studies appeared to be well conceived, but yielded somewhat confusing results due to apparent seasonal differences in the presence/distribution of individuals.

Brad Sample:

Although a site-specific mink toxicity study was conducted, the results of the TDI modeling were compared to literature-derived toxicity data. The site-specific mink study by Bursian et al. produced NOAELs and LOAELs for kit survival of 0.169 and 0.414 mg/kg/d for PCBs and 1.69 and 7.67 ng/kg/d for TEQs. In contrast, the literature-derived values were 0.128 and 0.0272 mg/kg/d for PCBs and 3.6 and 36 ng/kg/d for TEQ. The results from the site-specific toxicity study should be used to evaluate exposure and risk rather than the literature-derived data.

RESPONSE 3.6-BS-8

One of the fundamental principles of WOE analysis is that, to the extent possible, the analysis should rely on independent lines of evidence. Using an effect metric from the mink feeding study to compare to exposure estimates would result in non-independence of the feeding study and the comparison of exposure and effects lines of evidence. However, further discussion of the difference in effects data from the lines of evidence will be included in the revised ERA.

The site-specific mink study indicates that the Housatonic PCB mixture is somewhat less toxic to mink than mixtures observed at other locations. As a consequence, use of literature-based data over-estimates risks at the site. My recommendation is that the site-specific data be used instead if the literature-derived values.

RESPONSE 3.6-BS-9

EPA agrees that the Housatonic PCB mixture is less toxic to mink than observed at other sites. It should be noted that the site-specific data, rather than literature values, were used in the risk characterization to develop the MATC.

Ralph G. Stahl:

The effects metrics were reasonable and consistent, but were not objective. For one, there is no evidence to conclude that the numbers of mink or otter found in the PSA are depressed or linked quantitatively to the levels of tPCBs or TEQ in soils, sediments, surface water or biological tissues (fish or prey items). Thus to suggest there is a cause and effect is not objective.

RESPONSE 3.6-RS-4

As with most biological surveys, it is not possible to make a definitive link between the numbers of mink and river otter and the concentrations of contaminants in the Housatonic River watershed based on the EPA and GE field surveys. However, EPA believes that the evidence indicates that the numbers of mink and river otter are depressed in the PSA. Even if the cause of this decrease cannot be definitively attributed to tPCBs or TEQ, this information is an important component of the overall weight-of-evidence assessment for piscivorous mammals that is presented in the ERA.

Second, the feeding study in mink appears to have overstated the significance of the jaw lesions found in kits. While EPA argues that the lesion is potentially indicative of a pre-neoplastic process, there is no evidence presented to support this finding in kits allowed to reach adulthood. In fact, none of the kits appear to have been kept to an age sufficient to determine conclusively whether the jaw lesions lead to a malignant, potentially metastatic tumor or not. Relatedly, there is substantial speculation provided in the ERA indicating that those kits suffering from the jaw lesion may, over time, starve to death, and this too was suggested as a reason why the numbers of adult mink in the PSA appeared to be "low". Again, there are no data presented to support this finding in feral mink, since it does not appear that any adult feral mink were trapped or examined in the PSA. In the absence of such site-specific, confirmatory information, any conclusion regarding the jaw lesion leading to starvation and causing depression of the resident mink population in the Housatonic River watershed is purely speculative.

RESPONSE 3.6-RS-5

Please refer to the response to General Issue 15.B.

There is some debate as to the significance of the measurements conducted during the mink feeding study. On the one hand, the reproductive successes of the adult females do not appear to have been compromised from exposure to HR fish containing tPCBs and TEQ. Yet, kits born to the females did show a slight, but evident depression of body weight at 3 wks of age, and more significant depression in survival rate at 6 mos of age. While there are less than convincing data regarding impacts to reproductive health in the adult females, these results in the kits are of concern and were identified clearly in the ERA. I have identified another concern with the survival data discussed more fully in my later comments. This comment on the kit survival at 6 mos is germane to assessing whether the mink feeding study was conducted under the best of scientific practices.

RESPONSE 3.6-RS-6

EPA agrees with the Reviewer's comment that the results for kit survival at 6 weeks are significant; however, EPA notes that the study was not designed to evaluate reproductive health of adult females.

As stated above in Response 3.6-RS-1, all kits that died after 6 weeks of age were necropsied; the pathologist's assessment is contained in the mink feeding study report (all died of infections thought to be unrelated to PCBs). Kits that died prior to weaning at 6 weeks were not necropsied. EPA does not agree with the Reviewer's comments that the study may not have been conducted to best scientific practices because it was not possible to perform necropsies on kits less than 6 weeks of age. Previous mink studies have shown kit survivability to be a sensitive indicator of PCB toxicity in this species, and at least one study has demonstrated significant delivery of PCBs from female to kit during lactation (Bleavins et al. 1980; Heaton et al. 1995). Neither of these studies necropsied mink kits that died prior to 6 weeks of age. More importantly, deaths due to causes other than contaminants in the diet would be expected to randomly occur across dietary treatments.

Timothy Thompson:

The issue of jaw lesions should be further explored. While I concur that connecting the presence of jaw lesions leading to starvation is speculative, there is some evidence cited in the scientific literature in the ERA. It's also important to determine if the observed effects are hyperplasiac, or precancerous – the difference is important. Having said that, I do believe that the presence of these lesions can be supportive of risks in the PSA to mink.

RESPONSE 3.6-TT-4

Please refer to the response to General Issue 15.B.

3.6(e) Were the statistical techniques used clearly described, appropriate, and properly applied for the objectives of the analysis?

Valery Forbes:

The statistical methods in the mink feeding study do not seem to be described in the Methods section of this study (I.3.2.1.3). They should be briefly described here. See comments on modelling for wildlife in general.

RESPONSE 3.6-VF-3

Please refer to the response to General Issue 1.

Thomas W. La Point:

Yes. The feeding study was very well conducted and used appropriate controls. The results, well-quantified, are strong and telling.

James T. Oris:

See comments in 3.4. Much of the methodology is simply cited from other reports. The details need to be in the document.

RESPONSE 3.6-JO-6

EPA agrees with this comment. More methodological information will be provided in the revised ERA. Please refer also to the response to General Issue 1.

Mary Ann Ottinger:

The models for the mink and otter populations are reasonable, given the laboratory data and information from the literature. However, there may be an underestimate of the long-term impact to mink, based on the data collected. Although this was in addition to the some of the stated objectives, these data should be considered as relevant for long-term impact and potential population level impacts.

RESPONSE 3.6-MO-2

The effects metrics considered in the piscivorous mammals assessment (i.e., survival, growth and reproduction) are measures of chronic (long-term) impacts. Please also refer to the response to General Issue 9.

Brad Sample:

Although the statistical analyses in the mink toxicity study are well described and appropriate, the same transparency issues previously mentioned under other receptors apply for the analyses for mink and otter.

RESPONSE 3.6-BS-10

Please refer to the response to General Issue 1.

Similarly, the issues with the probabilistic analyses previously mentioned under other receptors apply for the analyses for mink and otter.

Ralph G. Stahl:

The statistical techniques were clearly described and appropriate.

Timothy Thompson:

The statistical methods were not provided in the sections. Same global comment for all of the sections on statistical methods.

RESPONSE 3.6-TT-5

Please refer to the response to General Issue 1.

3.6(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria?

Valery Forbes:

In general yes.

Thomas W. La Point:

Somewhat. The field surveys (either EPA or GE) have large variances in sampling and quantifying mink or otter use. This is completely understandable for such types of species as mink and otter. They are difficult to study in this manner. The primary results (feeding study and literature survey) should have very high weighting value – much higher than the field surveys. Given that, the risk rating should remain "high," not "intermediate to high."

RESPONSE 3.6-TL-3

The mink feeding study received the highest weighting possible in the weight-of-evidence assessment for piscivorous mammals (see Table I.4-4). The comparison of modeled exposure to effects line of evidence (i.e., the "literature survey") received a Moderate/High weighting. A high weighting was not given to this line of evidence because some of the modeling inputs were not site-specific and because of a few other minor limitations (see Table I.4-4). EPA does not believe that any changes are required for the feeding study or the comparison of modeled exposure to effects lines of evidence. However, as noted above, EPA will review the weights assigned to the field surveys in the revised ERA. Please refer also to the response to General Issue 8.

James T. Oris:

Combining mink and otter (Table I.4-4) in the risk characterization should not be done. They should be treated as separate receptors.

RESPONSE 3.6-JO-7

EPA agrees. Please refer to the response to General Issue 15.A.

Mary Ann Ottinger:

The risk characterization was supported by the literature and by the laboratory data. There were few field data to rely upon for the models.

Brad Sample:

Although the risk characterization was appropriate for the field studies, the results of the TDI modeling should be evaluated based on the site-specific toxicity test results and not on the literature-derived results.

RESPONSE 3.6-BS-11

Please refer to Response 3.6-BS-8 above.

Ralph G. Stahl:

The characterization of risk was not fully supported by the available information. In my opinion, the ERA is highly speculative with respect to the field observations and the reason for there being apparently "low" numbers of adult mink in the PSA. Finding and quantitatively documenting adult mink in the PSA and the watershed is a very difficult undertaking and the ability to conduct such a study is as dependent on proper implementation of field protocols as it is on a myriad of other confounding influences (weather, flood stage, food sources, human interferences, season, etc.)

RESPONSE 3.6-RS-7

EPA does not believe that the difficulties of conducting field surveys mean that the results from such surveys are "highly speculative". Well-developed protocols for surveying mink and river otter are available and were used by EPA. When followed properly, these methods are capable of detecting mink and river otter if they are resident or frequently forage in the area of interest. When field studies also include the use of reference areas, and are conducted at the same time, most confounding influences are controlled for.

Second, I have concerns about the discussion rendered for the jaw lesions, their etiology, and their potential for impacts on populations of mink exposed to tPCBs in the Housatonic River watershed.

RESPONSE 3.6-RS-8

Please refer to the response to General Issue 15.B.

The feeding studies provide the largest set of evidence of potential harm to kits and it is those data which appear to be the most appropriate under the evaluation criteria. Similar to results from the other assessment endpoints, there is a lack of strong concordance between the field and laboratory studies.

RESPONSE 3.6-RS-9

As discussed in Appendix I, Section 9, of the ERA, all three lines of evidence for piscivorous mammals suggest that tPCBs and TEQ pose risk to mink and river otter inhabiting the PSA of the Housatonic River.

Comments were provided by GE suggesting that the mortality of kits at 6 mos of age could be a result of causes other than exposure to tPCBs and TEQ. It appears that kits which died unexpectedly in the study were not necropsied to determine the cause of death. This is a legitimate point and should be given serious consideration in the revision to the ERA. It is standard practice in long term studies to determine the cause of death in any animal that dies unexpectedly during the treatment. Without this standard practice the many 2-yr carcinogenicity bioassays in rats would be of little value to assessing the carcinogenic potential of a substance. This is also true for the mink feeding studies and appears to be a serious oversight of the EPA contractors. If, on the other hand, the kits were necropsied and the cause of death ascertained, then that information is crucial to resolving the point raised by GE.

RESPONSE 3.6-RS-10

As stated above in Responses 3.6-RS-1 and 3.6-RS-6, all kits that died after 6 weeks of age were necropsied; the pathologist's assessment is contained in the mink feeding study report (all died of infections thought to be unrelated to PCBs). Kits that died prior to weaning at 6 weeks were not necropsied. EPA does not agree with the Reviewer's comments that the study may not have been conducted to best scientific practices because it was not possible to perform necropsies on kits less than 6 weeks of age. Previous mink studies have shown kit survivability to be a sensitive indicator of PCB toxicity in this species, and at least one study has demonstrated significant delivery of PCBs from female to kit during lactation (Bleavins et al. 1980; Heaton et al. 1995). Neither of these studies necropsied mink kits that died prior to 6 weeks of age. More importantly, deaths due to causes other than contaminants in the diet would be expected to randomly occur across dietary treatments. More detail in response to this comment is provided in Response 3.6-RS-1.

Timothy Thompson:

With the caveat of needing to look again at GE's comments on dose/response – this is a very well done study and supports the conclusions of the ERA.

RESPONSE 3.6-TT-6

Please refer to the response to General Issue 15.B.

3.6(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made.

Valery Forbes:

Yes.

Thomas W. La Point:

Yes, they were adequately addressed and acknowledged to be high. However, given the extensive literature data on the sensitivity of mink to PCB exposure, the uncertainty should be treated with conservatism in assessing risk to mink or otter. In my opinion, the only improvement would be to incorporate a radio-collar study of five to 10 mink and follow them over one or two years. The cost of such a study would be high – but relative to the value of the resource, the cost would be comparatively small.

RESPONSE 3.6-TL-4

EPA considered the use of telemetry methods when initially designing the field survey. However, the MA regulations restricting live trapping methods and the lack of observations during the initial ecological characterization suggested that such an effort (along with the required subsequent intensive monitoring) would be unlikely to result in enough, if any, animals to provide useable data sufficient to justify the level of effort.

Also, GE proposed a mink radiotelemetry study but abandoned the effort when they could not locate any mink in the PSA prior to winter tracking efforts.

- On July 25, 2002, GE presented a PowerPoint presentation to EPA outlining their ecological studies. On slide 45, GE described that the mink study had three components, the first of which was: "Radio-telemetry/mink trapping study, Live –Trapping attempted in spring of 2001 – no mink caught so component dropped."

- In the *Evaluation of Mink – Presence/Absence, Distribution, and Abundance in the Housatonic River Floodplain* report (BBL, September 2002), on page 1, third paragraph, last sentence, it is stated that "In addition, we attempted to trap mink for a radio-telemetry study' however, no mink were captured and this portion of the study was discontinued."

James T. Oris:

Uncertainties were appropriately described. Analytical uncertainties should be included.

Mary Ann Ottinger:

The uncertainties resent in the field studies were discussed. As discussed for other receptors, the traditional measurement end points are likely to be inadequate to detect more subtle sub lethal effects of some toxins, such as endocrine disrupting chemicals. The lesions observed in the mink may be a symptom of immune system impact. Although not used in these studies as a measurement end point, future assessments may be more comprehensive if some more subtle measurement end points are included.

Brad Sample:

Although the uncertainties are generally well captured in the ERA report, the relative importance of various uncertainties is not discussed. Interpretation and understanding of the results would benefit by such an analysis. Spatial heterogeneity of contamination and the degree to which the exposure data actually represent the PSA are very significant issues. These need to be discussed.

RESPONSE 3.6-BS-12

EPA agrees with this comment, and a discussion of the relative importance of the sources of uncertainty will be provided in the revised ERA.

Ralph G. Stahl:

The significant uncertainties in the analysis of the assessment endpoints were identified and adequately addressed.

Timothy Thompson:

The uncertainty in this section was well characterized. Same global comments regarding analytical variability are applicable here.

3.6(h) Was the weight of evidence analysis appropriate under the evaluation criteria? If not, how could it be improved?

Valery Forbes:

The WOE table should be separated for mink and otter because different kinds and amounts of information were available for each species.

RESPONSE 3.6-VF-4

EPA agrees. Please refer to the response to General Issue 15.A.

Thomas W. La Point:

The WOE analysis "overweighted" the field study, in my opinion. There is sufficient literature and, when included with the mink feeding study, the weighting on the modeling efforts should be increased, leading to a decision of "high risk" for Piscivorous mammals feeding within the PSA.

RESPONSE 3.6-TL-5

EPA will review the weights assigned to the field surveys. Please refer to the response to General Issue 8.

James T. Oris:

See previous comments on WOE analysis. GE studies should be downweighted due to methodological limitations.

RESPONSE 3.6-JO-8

EPA will review the weights assigned to the field surveys. Please refer to the response to General Issue 8.

Mary Ann Ottinger:

The WOE was appropriate given the data from lab, literature, and to a lesser extent, field data. The same comment as in 3.7g is pertinent.

Brad Sample:

WoE analyses had a couple of inconsistencies that need to be corrected.

- The WoE analyses combined both mink and otter. This is inconsistent with other sections. Separate WoE evaluations need to be performed for mink and otter.

RESPONSE 3.6-BS-13

EPA agrees. Please refer to the response to General Issue 15.A.

- TEQs were listed as an independent line of evidence – this is not correct and is inconsistent with other sections.

RESPONSE 3.6-BS-14

EPA agrees that there are some inconsistencies between how tPCB and TEQ assessments were presented in the risk characterization section for piscivorous mammals compared to, for example, omnivorous/carnivorous mammals. In the ERA, the TEQ and tPCB assessments were conducted separately, as they were in all of the wildlife assessments. However, the presentation format used in Table I.4-4 may have given the incorrect impression that the assessment for piscivorous mammals had four independent lines of evidence (i.e., field survey, feeding study, modeled exposure, and effects for tPCBs and for TEQ). The presentation format will be corrected in the revised ERA to reduce the potential for confusion.

Ralph G. Stahl:

The weight of evidence analysis was appropriate under the evaluation criteria. As noted earlier, the speculation in the ERA on jaw lesions/starvation being the cause of "low" numbers of adult mink in the PSA should be excised.

RESPONSE 3.6-RS-11

Please refer to the response to General Issue 15.B.

Timothy Thompson:

See previous comments.

3.6(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?

Valery Forbes:

Yes, these are ok.

Thomas W. La Point:

Please refer to my responses in 3.6.g and 3.6.h. In my opinion, the risk estimates as provided in this ERA (= "moderate to high") were derived by providing too much weight to the field studies.

RESPONSE 3.6-TL-6

EPA will review the weights assigned to the field surveys. Please refer to the response to General Issue 8.

James T. Oris:

Yes.

Mary Ann Ottinger:

Yes, based on the measured tPCBs downstream.

Brad Sample:

The downstream risk evaluation for mink and otter appears to be appropriate. However there are several issues that need to be addressed before a definitive conclusion can be made:

- The downstream analyses is only a screening evaluation and significant uncertainties exist. Text needs to be added clarifying that the results only represent a screen and the uncertainties specifically associated with this evaluation need to be identified.

RESPONSE 3.6-BS-15

EPA agrees with the Reviewer's comment that there are greater uncertainties associated with the downstream assessment, but does not agree that it is only a screening analysis. EPA will include a more comprehensive discussion of sources of uncertainty identified in the downstream assessment in the revised ERA.

- The analysis lacks transparency. The text briefly describes how the MATC was derived and states that exposure was based on fish data collected in the downstream reaches between 1999 and 2002. Only in Figures I.4-15 and I.4-16 is it stated that manipulations of the fish concentration data were conducted to estimate exposure. It appears that various scaling factors were employed to adjust fish concentration data. A detailed explanation and justification of these adjustments needs to be provided in the text. The sources for the scaling factors applied needs to be presented.

RESPONSE 3.6-BS-16

EPA agrees that more detail explaining how the MATC was derived and how some of the fish concentrations were extrapolated would improve the clarity of this discussion. This information will be provided in the revised ERA.

- Risks are broadly displayed in Figures I.4-15 and I.4-16 as reaches of the river where exposure either exceeded the MATC or did not. Locations from where fish samples are available are not presented. I suspect that the available data are not spatially continuous enough to allow such broad conclusions. Locations from where fish samples exist need to be presented in the map and the risk conclusions should be based solely on those areas for which fish data are available.

RESPONSE 3.6-BS-17

Please refer to the response to General Issue 1. EPA does not agree with the comment that the downstream assessment can only be conducted "solely [for] those areas [where] fish data are available". The samples are considered representative of the reaches from which they were collected and thus the risk conclusions based on these samples can be reasonably extrapolated to the entire reach.

- What contaminants were measured in the downstream fish?. Were the analyses restricted to PCBs? As the analyses for Reaches 5 and 6 were based on PCBs and TEQ, for consistency, a statement as to whether TEQ data were available downstream needs to be made. If TEQ data are available, they should be evaluated also.

RESPONSE 3.6-BS-18

Few fish and aquatic invertebrate samples downstream of Woods Pond have the congener-specific analyses required to calculate TEQ as these analyses were not required in the historical or ongoing long-term monitoring programs. Thus, it was not feasible to extend the TEQ assessment downstream of Woods Pond for wildlife that prey on aquatic biota.

- The upstream and downstream exposure analyses differed in how they were conducted – (more complete dietary exposure estimation for Reaches 5 and 6, screening of MATC against fish concentrations in downstream reaches). For the purpose of comparing relative exposures and risks along the course of the river, additional screening of the fish data (using the MATC approach used for downstream locations) from reaches 5 and 6 should be considered. These data could be presented in this part of the section as histograms representing each reach.

RESPONSE 3.6-BS-19

EPA agrees that a graphic presenting fish concentrations upstream and downstream of Woods Pond would provide a useful relative comparison of risk to piscivorous wildlife. This information will be provided in the revised ERA.

Ralph G. Stahl:

The risk estimates were objectively and appropriately derived for reaches of the river where site-specific studies were not conducted.

Timothy Thompson:

Same criteria from Assessment Question 1.

3.6(j) In the Panel members' opinions, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?

Valery Forbes:

The ERA concluded that the risk to piscivorous mammals is high and that confidence in this conclusion is high. I believe the magnitude of the risk to piscivorous mammals is more uncertain than indicated due to the limited site-specific information on which it is based.

RESPONSE 3.6-VF-5

The conclusion of risk to piscivorous mammals is based, in large part, on substantial site-specific information. Two site-specific biological surveys were conducted for piscivorous mammals. In addition, the fish used in the mink feeding study were collected from the PSA, and site-specific prey concentration data were used to model exposure of mink and river otter to tPCBs and TEQ.

Thomas W. La Point:

No, not strictly. I think the evidence points towards a greater risk for piscivorous mammals than is presented in this ERA.

RESPONSE 3.6-TL-7

EPA concluded that the risk to piscivorous mammals was high, but in the revised ERA, EPA will review the description provided in the risk assessment conclusion.

James T. Oris:

Yes, supported.

Mary Ann Ottinger:

Yes, based on the models and information presented.

Brad Sample:

I believe that the risk conclusions are probably supported by the available data. Revisions of the analyses as suggested above are necessary however to ensure that the appropriate magnitude of risks is documented.

In addition, some revision and clarification is needed to indicate that the evaluation is for the local population and not for individuals.

RESPONSE 3.6-BS-20

The conclusions regarding risks of tPCBs and TEQ to piscivorous mammals are for the locally exposed population. Please refer also to the response to General Issue 9.

Ralph G. Stahl:

In my opinion, the information provided in the ERA and the evaluation thereof generally support the conclusions regarding intermediate risk to local populations of mink. The feeding studies are not without problems, but are the most robust of the data evaluated that suggest the potential for high risk to kits. I temper this statement with my concerns on whether the kits that died unexpectedly were necropsied and the cause of death determined. On this point, the EPA should give serious consideration to re-examining any archived whole bodies or organs from the kits that died unexpectedly.

RESPONSE 3.6-RS-12

Please refer to Responses 3.6-RS-1, 3.6-RS-6, and 3.6-RS-10.

Further, as noted in the ERA, there was no definitive dose-response between kit survival and tPCB content of the fish, which further clouds the conclusion that might be drawn from the study.

RESPONSE 3.6-RS-13

EPA believes that there was a definitive dose-response relationship demonstrated for survival of mink to six weeks of age in the mink feeding study. Please refer to the response to General Issue 15.B.

Even so, results from other mink feeding studies in the Great Lakes have clearly demonstrated potentially harmful effects in mink which cannot be overlooked in the context of mink consuming tPCB-contaminated fish within the Housatonic River watershed. However, the magnitude of the risk to mink posed by exposure to contaminated fish in the watershed appears to be less than that reported from studies in the Great Lakes and elsewhere. This also supports the categorization of risk to mink as intermediate.

RESPONSE 3.6-RS-14

EPA believes that the risk to piscivorous mammals is high, as characterized in the risk assessment conclusion in the revised ERA, even for those animals that forage in the PSA only a small portion of the time.

Timothy Thompson:

Yes.

3.7 Omnivorous and Carnivorous Mammals

> ***3.7.(a) Were the EPA studies and analyses performed (e.g., field studies, site-specific toxicity studies, comparison of exposure and effects) appropriate under the evaluation criteria, and based on accepted scientific practices?***

Valery Forbes:

Yes as far as I can determine.

Thomas W. La Point:

The strongest studies in this component were the modeling studies, using literature (rat) data and the GE demographic field study (with re-analysis). However, all studies indicate an extremely large degree of uncertainty, which precludes any strong statement of harm or risk. However, that said, the studies that were conducted did follow accepted scientific procedures. The semi-quantitative study of shrews collected in the amphibian study are weak – and should lead to a conclusion that a study focused on omnivorous field mice, opossums, or raccoons (using radio collars) would have been beneficial. Yet, that is hindsight…..

James T. Oris:

Generally appropriate. I question the combination of all of the possible receptors in this feeding category into one receptor. Justification of this combination needs to be made. For example, is it appropriate to combine shrews and fox? It may not be reasonable to separate out all of the different feeding strategies and guilds, but some discussion of how this may add to uncertainty and how the outcome of the risk assessment may be affected is warranted.

RESPONSE 3.7-JO-1

EPA agrees. Please refer to the response to General Issue 16.A.

Mary Ann Ottinger:

Four years of field surveys showed species diversity in the PSA, with some indication of differential distribution of the species. Representative species chosen were red fox and northern short-tailed shrew. Foxes were observed throughout the PSA; however the models relied on data from the literature. Female mammals (primarily white-footed mice) were trapped; counting placental scars assessed fecundity. This is a good measure, but placenta scars in adult females should be considered in combination with estimates of survival of the young of that year, total number of animals trapped, sex ratio, and percentage of females that were immature or infertile.

RESPONSE 3.7-MO-1

EPA considered many of these metrics in the assessment for short-tailed shrews, particularly with the shrew field study line of evidence (please see Appendix J, Section 4.3.3).

Brad Sample:

The EPA studies appear to be suitable and appropriate.

Ralph G. Stahl:

The EPA field studies were performed appropriately under the evaluation criteria and based on accepted scientific practices. There were no site-specific toxicity studies conducted for this assessment endpoint.

I do not believe the estimate of effects for red fox was appropriate under the evaluation criteria. The toxicity reference value used for the comparison of exposure and effects was based on rodent data. Rodents are physiologically, substantially dissimilar from red fox. This makes any comparison highly uncertain and, in my opinion, not useful in the context of this ERA.

RESPONSE 3.7-RS-1

EPA's review of the effects literature for red fox did not find any canine toxicity data for tPCBs or TEQ. EPA, therefore, relied on effects data for rodents and the uncertainty of this extrapolation is clearly acknowledged in Section J.4.5. This extrapolation is similar to what is routinely done in human health risk assessment (rodent to humans), and is necessary for the vast majority of wildlife assessments. EPA will update the literature review to look for recent canine effects data, and if new data are found, will revise the ERA for fox. Otherwise, EPA will continue to rely on rodent effects data to estimate risk of tPCBs and TEQ to fox.

Timothy Thompson:

Studies appear to be appropriate for the risk assessment and are based on accepted scientific practices.

3.7(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria, based on accepted scientific practices, and incorporated, appropriately in the ERA?

Valery Forbes:

Yes, the demographic study was a useful contribution (despite the sampling problems apparently corrected by EPA's reanalysis (or not)). GE's reanalysis of EPA's reanalysis using the same data but a slightly different statistical technique gave a non-significant result. GE treated the grids as replicates and did not weight them differentially, as a function of sample size within grids, used probit transformation, and did not include grid 3 in the analysis of males because there was only one male. In contrast, the EPA used the method of Baylor & Oris (1997; ET &C), probit transformation, and weighted grids according to the total number of organisms in the treatment. Given the dependence of the statistical significance on subtle differences between two (seemingly) appropriate statistical methods, the most robust conclusion that can be made from this study is that the response is borderline.

RESPONSE 3.7-VF-1

EPA will include a discussion of both statistical approaches in the revised ERA. Please refer to the response to General Issue 16.B.

Thomas W. La Point:

With the re-analysis of original results, the GE short-tailed shrew study was very strong and indicates a high level of potential harm (at least in Location 13) to insectivorous small mammals. Given the well-published consequences of PCB exposure on mortality and reproductive effects in mammals, the results from this study that relate soil PCB concentration to mortality (albeit with broad confidence limits) should carry high weight.

Response 3.7-TL-1

Using appropriate statistical methods, EPA demonstrated a significant dose-response relationship between PCB concentrations in soil and shrew mortality. In response to the Reviewer's comments regarding weight-of-evidence, EPA will re-evaluate the weight assigned to this field study. Please refer to the response to General Issue 16.B, as well as the preceding specific response, for further discussion of this issue.

James T. Oris:

As with all GE studies, this one was designed to address a very narrow question and suffered from design and analysis limitations. There were no reference sites, sample areas were limited in number and had high variation in habitat quality. Habitat quality was not addressed. The data were analyzed using each grid as a single replicate, thus there was a sample size of 3 "high" PCB and 3 "low" PCB sites. This is minimal sample size and thus renders statistical analysis nearly powerless. The lack of power allows for the analyst to conclude "no effect" when, in fact, there may be an effect. The study was designed to support a "no effect" conclusion and thus should be severely discounted.

Upon reanalysis, the EPA showed a significant relationship between PCB concentration and shrew survival. The two results are so divergent that it is clear that GE and EPA conducted very different analyses. Upon question at the public panel meeting, the differences were clarified.

In my opinion, the GE analysis was not appropriate for this particular data set. The use of standard probit analysis is not supported by the data quality. GE and EPA techniques are equivalent only if there is a montonic dose response relationship. The probit analysis used by GE cannot account for curvature in data. However, the Bailer and Oris technique used by EPA can model dose responses that include curvature. When a dose response relationship is nonlinear, this will result in different answers between the two methods.

In the case of the shrew data, the dose response was non linear. Use of standard probit model in this case is inappropriate because it violates requirements for a valid probit (e.g., don't use "zero effect" doses, need 3 partial effect levels, monotonic (linear) response. The Bailer&Oris method can be used to parameterize for any shape of dose response relationship, using a polynomial regression and specifying the appropriate link function (e.g., probit or logit transformation) and error distribution (e.g., binomial for survival or presence/absence). The Bailer&Oris method with a 1st order polynomial (i.e.., a straight line), a probit link, and a binomial error distribution is equivalent to a standard probit regression. The Bailer&Oris method, however, with a 2nd order polynomial will fit a smooth. In this case, the straight line probit would indicate that there was no significant effect within the concentration range tested, but the Bailer&Oris model would.

As a separate issue, the GE analysis used only one data point at each site. If this is the mean of several points, you have thrown-out data and reduced the power of testing for differences between concentrations. The use of individual data points instead of using the means will generate tighter confidence limits around the regression line, the slope value, and the intercept value. This could make the difference between saying that the slope of the line is not different from zero (i.e.., no dose response) versus saying that the slope is different from zero (i.e.., yes, dose response). In other words, using just mean values in the regression could result in an incorrect conclusion of no dose response when one, in fact, existed, regardless of the shape of the dose response (straight line or curved). These two issues (curved dose response and using the mean value at each concentration) compound the situation and make it even easier to incorrectly state that there is no dose response.

Response 3.7-JO-2

EPA agrees with the comments made by the Reviewer regarding some of the limitations of the shrew field study (i.e., small sample size results in low statistical power, study lacked reference sites, and there was high variation in habitat quality between sites). For a detailed discussion of the issues related to the GE statistical analysis and the re-analysis conducted by EPA, please refer to the response to General Issue 16.B.

Mary Ann Ottinger:

The Boonstra study (2002) showed no negative effects of tPCBs on short-tailed shrews. There are some difficulties in interpretation of these data without a population assessment; stability in population may be due to immigration.

Response 3.7-MO-2

The statistical re-analysis conducted by EPA indicated a significant relationship between soil concentrations of tPCBs and shrew survival, although the confidence limits indicated that the relationships were not strong. Thus, EPA does not believe that the Boonstra study was able to demonstrate “no negative effects of tPCBs on short-tailed shrews.” EPA agrees with the comment from the Reviewer that immigration or other factors may be contributing to the relatively high numbers of shrews observed in areas with high tPCB concentrations in soil.

Brad Sample:

This short-tailed shrew study was suitable and used appropriate methods. However, there are a number of study design issues that limit the utility of its results. The study suffers from the lack of a reference location, small sample sizes, and limited study duration. Habitat quality also appears to be a confounding factor that was not adequately evaluated. Because the data were from only a single year - the representativeness of conditions in the year are unknown. No data from a reference locations were available for comparison. Sample sizes were very small. Vegetation cover was heterogeneous among and within sampling areas. Statistics showed that population parameters differed significantly by area - strongly suggesting a confounding habitat effect - Addition analyses using detailed habitat data may be useful to tease out the habitat effects from contaminant effects. Given the high variability known to exist for shrew populations over years, seasons, and habitats, etc., the fact that the study looked only at one year, had no reference area, and had only 6 locations seriously limits the strength of any conclusions.

RESPONSE 3.7-BS-1

EPA agrees with the comment regarding the possible influences of habitat on the study results. If the data permit, EPA will conduct additional analyses to determine whether habitat-specific characteristics could have obscured effects of tPCBs and TEQ on short-tailed shrews in the field study, and will include the results in the revised ERA. The other limitations of the shrew field study were considered in weighting this particular line of evidence (see ERA Table J.4-7).

Conflicting conclusions based on the GE and EPA analyses of data from this study need to be better explained. At the January meeting, GE stated that they could not recreate the statistical results produced by EPA. Extensive discussion at the meeting indicated that the analyses conducted by EPA and the re-analyses conducted by GE differed in how they handled the data and the exact statistical methods used:

GE/Boonstra – treated grids as reps - did not weight grids – did probit transformation – eliminated grid 3 due to presence of single male – not included in analyses for males, included for male and female.

EPA – methodology in Bailor and Oris – used GliM models with – probit transformed link function – each grid is weighted by total number of organisms in the treatment.

I cannot comment on the statistical applicability of both approaches. I do think that the Boonstra re-analyses should be included in the ERA and the differences between the two approaches and their results should be discussed.

RESPONSE 3.7-BS-2

EPA will include a discussion of both approaches in the revised ERA, please refer to the response to General Issue16.B.

In the second paragraph of Section J.4.3.3 it is stated that the spatially weighted PCB concentrations were calculated based on the *arithmetic mean* concentration in soils. For virtually all other analyses in this ERA, the lognormal distribution was assumed for abiotic media. An explanation/justification for the use of the arithmetic mean needs to be provided.

RESPONSE 3.7-BS-3

Short-tailed shrews forage over distances of tens of meters and more over a period of weeks to months. Thus, individuals tend to integrate spatial variation in the concentrations of soil and prey to which they are exposed. In the absence of significant metabolism or excretion of a contaminant, average exposure will be approximately equal to the sum of exposures over time divided by the length of exposure (i.e., the arithmetic mean). Thus, even though soil concentrations likely have an underlying lognormal distribution, the arithmetic mean is the appropriate metric to account for spatial and temporal averaging of exposure to COCs in soil or prey by shrews.

Ralph G. Stahl:

The Boonstra study conducted in small mammals was appropriate under the evaluation criteria, but was not without design problems. The study duration was limited and there does not appear to be any reference areas evaluated in this study design. Thus it may not have been based on accepted scientific practices. Nevertheless, the study was incorporated appropriately into the ERA.

Timothy Thompson:

This is not my area of expertise. Having said that, the study appears to be well done, and follows solid scientific practice and principle. The largest outstanding issue that should be addressed is the apparent discrepancy between GE's analysis, and the re-analysis conducted by EPA. GE's reanalysis of EPA's reanalysis using same data and technique gave a non-significant result. GE/Dr. Boonstra treated the grids as replicates and did not weight them differentially, as a function of sample size within grids. They used probit transformation following the methods of Sokal and Rolf (pp 544-547). Grid 3 was not included in the analysis of males because there was only one male. EPA followed the method of Baylor & Oris (1997; ET &C); general linear modeling with probit transformation. The grids were weighted according to total number of organisms in the treatment.

Summarizing from Dr. Samples notes (which I concur with), while the analyses are different, they are both likely valid. Both are probit regressions; simple probit (Boonstra rework) is a monotonic linear model, while the Bailor and Oris model can account for non-linear response and account for error bounds differently. Bottom line is that the results indicate that response is not strong – both methods should probably be reported and discussed in final ERA.

RESPONSE 3.7-TT-1

EPA will provide a discussion of both methods in the revised ERA; please refer to the response to General Issue 16.B.

3.7(c) Were the estimates of exposure appropriate under the evaluation criteria, and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?

Valery Forbes:

Yes.

Thomas W. La Point:

Yes, the spatially-weighted estimates for shrews was particularly appropriate, as shrews have such limited daily ranges. The estimates for the red fox are much less quantitative and have huge uncertainties associated. However, the uncertainties were acknowledged in the ERA. In my opinion, the re-analysis of the GE shrew demographic study provides some excellent trend data towards demonstrating consequences of exposure to soil-bound PCBs.

James T. Oris:

Within the limitations of using literature-derived values, the exposure estimates are appropriate.

Mary Ann Ottinger:

Field studies attempted to assess population and fecundity and relate these data to tissue residues from trapped animals. Both types of data are important. Additional data from a lab study with mice exposed to contaminated soil would have been valuable. This type of a study would have provided more direct information about the response of the species and the viability of offspring. It is possible that there has been some selection of a subset of the population that could survive and reproduce in the PSA environment with the COCs.

RESPONSE 3.7-MO-3

EPA agrees that additional field studies with rodent species would improve the ERA for omnivorous and carnivorous mammals. Due to practical limitations, however, it is not possible to conduct such studies.

Brad Sample:

Although the estimation of exposure and effects was generally appropriate, there were several areas that require further evaluation:

The TDI modeling effort for fox and shrews relied exclusively on directly measured concentrations of PCBs and TEQ in earthworms. This approach will give good estimates for all locations from which these site-specific biota data were available. However it does not address exposure and risk in areas from which site-specific biota data were unavailable, nor does it allow exposure and risk levels to be associated with specific PCB and TEQ concentrations in floodplain soils or sediments (information extremely useful to make remedial decisions). No effort was made to model bioaccumulation into biota and estimate exposure at all locations, with suitable, species-specific habitat, from which abiotic data were available. Given the large size of the site, the heterogeneous nature of contamination, and the variable movement patterns of resident wildlife, the current approach may not adequately represent the exposure that local populations of wildlife may experience at the site. It is recognized that modeled concentrations in biota increase uncertainty. However, this would be balanced by the increased spatial resolution in exposure estimates. I do not recommend replacing the current approach. Rather, exposure based on COPC concentrations modeled in biota should be used to provide a spatially more complete estimate. The results from both can be compared as part of the risk characterization.

RESPONSE 3.7-BS-4

EPA agrees that the risk characterization should include a broader assessment for areas of the PSA that were not directly included in the study design. In the revised ERA, EPA will either: (1) compare soil concentrations at locations where prey data were available (e.g., earthworm and terrestrial invertebrate sampling locations) to soil concentrations from other areas of the floodplain, or (2) develop BSAF model to enable estimates of exposure across the PSA. Either approach should provide a better characterization of risk across the entire PSA for those wildlife species that had risk estimates at only a few selected locations in the current ERA.

The TDI modeling effort relied exclusively on directly measured concentrations of COPCs in dietary items (either invertebrates for shrews, or small mammals for the fox). The results of these analyses, derived from data representing a comparatively limited spatial area, were then extrapolated to represent the whole PSA.(This is particularly an issue for the shrew – PSA exposure is based on invertebrate concentrations from only three locations. For the fox, the issue is that it is not clear from where the small mammals samples were collected – are they representative of the whole PSA?

RESPONSE 3.7-BS-5

The EPA small mammal locations were selected to be representative of the range of concentrations in the PSA. EPA agrees, however, that the risk characterization should include a broader assessment for areas of the PSA that were not directly included in the study design and will include such an assessment in the revised ERA (see preceding response). EPA also agrees that it is not clear from Appendix J where the small mammal samples were collected for estimating exposure to red fox (however, the locations were provided in Appendix A.1). EPA will describe the locations in Appendix J of the revised ERA.

Do they adequately represent the range of PCB and TEQ concentrations in soils in the PSA? Details on small mammal sampling locations needs to be included in this section.) For this approach to be valid, it is important that the areas from which biota data were collected represent the range of possible COPC concentrations in abiotic media, in particular it is essential that the highest concentrations be represented. It is not apparent that any analyses were conducted to indicate that the available site-specific biota adequately represented the full range of possible COPC exposures (including maximum concentrations).

RESPONSE 3.7-BS-6

Please refer to Response 3.7-BS-4.

References for sources of parameter values in Tables J.2-2 and J.2-3 and Tables J.2-10 and J.2-11 are lacking and should be provided.

RESPONSE 3.7-BS-7

EPA agrees and will include these references in the revised ERA.

Ralph G. Stahl:

The estimates of exposure were appropriate under the evaluation criteria.

Timothy Thompson:

See comment above.

3.7(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria?

Valery Forbes:

These were based largely on literature review and used largely rat and mouse data. Though not ideal this is probably acceptable.

Thomas W. La Point:

Yes. The emphasis on reproductive effects is good. Given the quantified mortality in the GE shrew study, and with the literature on rats indicating substantial reproductive effects at low-level, chronic exposure to PCBs, this metric is valuable.

James T. Oris:

Yes, appropriate.

Mary Ann Ottinger:

The effects metrics were appropriate with some difficulties in interpretation.

Brad Sample:

The effects data appear to be suitable and appropriate. EPA has conducted a very thorough review of the available toxicity data and has made a very strong effort to adequately represent the range of potential effects.

All toxicity data were based on rodents. These data may not taxonomically be the most suitable for fox. As they are from the same order (Carnivora), mustelid data may be more suitable for fox than rodent data. At a minimum, the potential effects of using data from taxonomically dissimilar species should be included in the uncertainty section (existing text expanded).

RESPONSE 3.7-BS-8

Surrogate species for red fox should ideally be closely related with regard to taxonomy, diet, physiology, and other factors. The red fox is omnivorous, feeding primarily on mammals including mice and rabbits, along with birds, and when available, fruits. Although rodents are taxonomically more distant from red fox than are mustelids, the rodent species for which toxicity data were extrapolated to red fox (i.e., Norway rat) are omnivorous, terrestrial, and have a diet that overlaps the diet of red fox. The mustelids for which toxicity data for tPCBs and TEQ are available (i.e., mink and, to a lesser extent, river otter) are piscivorous for the most part (they also consume other aquatic prey), aquatic or semi-aquatic, and have a diet that has little overlap with the diet of red fox. It is for these reasons that EPA chose to use toxicity data for Norway rats to extrapolate to red fox rather than use mustelid toxicity data. A discussion of this topic will be included in the revised ERA.

In the last paragraph of Section J.3.2.1 it is stated that the mouse was used as the surrogate for the shrew and the rat was the surrogate for the fox. It is unclear as to why this distinction was made. I'm assuming it was due to body size of the surrogate relative to the shrew or fox. The reasoning for this distinction should be explained.

RESPONSE 3.7-BS-9

The decision to use the mouse as a surrogate for short-tailed shrew, and the rat as a surrogate for fox was based primarily on body size considerations, as suggested by the Reviewer. In the revised ERA, EPA will include an explanation of why this distinction was made.

Ralph G. Stahl:

The effects metrics were appropriate under the evaluation criteria. However, the use of rodent data to estimate potential effects in the red fox is not, in my opinion, scientifically defensible.

RESPONSE 3.7-RS-2

Please refer to Response 3.7-BS-8.

Timothy Thompson:

The literature summary is impressive and appears substantively complete. The only comment I had was that there did not appear to be any canine studies with PCBs; I found that surprising.

RESPONSE 3.7-TT-2

EPA was unable to locate any canine toxicity data for tPCBs or TEQ in the review of the effects literature performed while developing the effects data set for the ERA. EPA will update the literature review to determine if recent canine studies have been published, and any such studies will be included in the revised ERA.

3.7(e) Were the statistical techniques used clearly described, appropriate, and properly applied for the objectives of the analysis?

Valery Forbes:

Yes.

Thomas W. La Point:

Including the re-analysis of the GE shrew survey, yes.

James T. Oris:

See comments in 3.4

Mary Ann Ottinger:

These models are also primarily based on the literature. The results for the EPA and GE data sets should be compared and aligned. If the difference is due to omission of data, then the rationale for leaving some of the data out of the analysis must be considered.

RESPONSE 3.7-MO-4

EPA agrees. Please refer to the response to General Issue 16.B.

Brad Sample:

Similar issue with transparency of statistical analyses as discussed for earlier sections, apply to this section too.

RESPONSE 3.7-BS-10

Please refer to the response to General Issue 1.

Similarly, the issues with the probabilistic analyses previously mentioned under other receptors apply for the analyses for shrew and fox.

Ralph G. Stahl:

The statistical techniques were not clearly described, nor based on the January 2004 public meetings, applied properly for the objectives of the analysis. The debate between EPA and GE concerning the proper statistical evaluation of the Boostra study leaves the reader and this panel member confused and frustrated. I have recommended that EPA and GE settle on one mutually agreeable approach and use it in the ERA.

RESPONSE 3.7-RS-3

EPA will include a description of both approaches in the revised ERA. Please refer to the response to General Issue 16.B.

Timothy Thompson:

Previous comments regarding transparency relevant here.

3.7(f) *Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria?*

Valery Forbes:

This appears appropriate.

Thomas W. La Point:

Results from modeling exposure and effects (largely using rat data) suggest that PCBs in soils pose a high risk to short-tailed shrews inhabiting Locations 13 and 14. In my opinion, much less can be said of the "intermediate risk" to red foxes exposed to tPCBs and TEQ in the PSA. The large uncertainty concerning the modeling line of evidence for tPCBs and TEQ forces the conclusion that risks of these COCs cannot be really estimated for the PSA. The GE shrew survey (with re-analysis) indicates that the small, insectivorous and/or omnivorous mammals are at risk. However, there may be extensive immigration into the PSA. This is simply unknown at present.

RESPONSE 3.7-TL-2

EPA agrees that there is high uncertainty associated with the conclusion of risk to red fox. The sources of uncertainty were discussed in Section J.4.5 of the ERA.

James T. Oris:

Yes, appropriate.

Mary Ann Ottinger:

The overall risk characterization is supported by the information, but there is uncertainty due to the lack of effects found in the field study in shrews. Potentially, there should be separate risk assessments for some of the species considered, especially those with greatly differing life histories.

RESPONSE 3.7-MO-5

EPA agrees that the discussion of the two species should be separated and will do so in the revised ERA. Also, please refer to the response to General Issue16.A.

Brad Sample:

The risk characterization is suitable and appropriate.

Organizational note: descriptions of the field studies and their analyses only appear in the Risk Characterization section. As these are really effects data, they should be described in the Effects section, with interpretation and integration of results left in the Risk Characterization section.

This anomalous presentation of data also occurs for some other receptors. The organization of the report would benefit if the presentation of the data were moved.

RESPONSE 3.7-BS-11

EPA agrees that the field studies should be moved to the effects characterization and will do so in the revised ERA.

Ralph G. Stahl:

The characterization of risk to small mammals is supported by the available information and appropriate under the evaluation criteria. However, the assignment of risk may change if the statistical analysis can be conducted to the mutual agreement of EPA and GE. Until that time, there will continue to be some concern as to whether the characterization of risk to small mammals is overly conservative or not.

RESPONSE 3.7-RS-4

Please refer to the response to General Issue 16.B.

The characterization of risk to red fox is not supported by the available information nor appropriate under the evaluation criteria. The modeled effects are based entirely on a comparison to rodent data which is not scientifically defensible in my opinion. I therefore recommend that the red fox assessment endpoint be deleted from the ERA evaluation. It provides little useful information, and will likely have no influence on the risk management decision.

RESPONSE 3.7-RS-5

EPA agrees that there is high uncertainty associated with the conclusion of risk to red fox because only one line of evidence is available to assess risk. EPA acknowledged this uncertainty in Section J.4.5. EPA does not, however, concur that the red fox assessment should be removed from the ERA, as this species is currently the only widely foraging, omnivorous and carnivorous mammal included in the ERA.

Timothy Thompson:

Yes.

3.7(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made.

Valery Forbes:

An improvement would be more site-specific information on both exposure and effects of mammals.

Thomas W. La Point:

There is so much uncertainty in this component (insectivorous and carnivorous small mammals) that conservatism must be called for, unless further emphasis is placed on quantifying the distribution and behavior of red foxes and other small mammals (raccoons, opossums, and the abundant number of mouse species). In this analysis, as with the Piscivorous mammals, above, a detailed study of red foxes, with radio collars, would be invaluable.

RESPONSE 3.7-TL-3

EPA agrees that while additional field studies would further reduce uncertainty in the ERA, practical limitations preclude the conduct of any further studies at this time.

James T. Oris:

Uncertainties were appropriately described.

Mary Ann Ottinger:

There were some inconsistencies in GE field study, leading to a reanalysis of the data by EPA.

Brad Sample:

Although the uncertainties are generally well captured in the ERA report, the relative importance of various uncertainties is not discussed. Interpretation and understanding of the results would benefit by such an analysis. Spatial heterogeneity of contamination and the degree to which the exposure data actually represent the PSA are very significant issues. These need to be discussed.

RESPONSE 3.7-BS-12

EPA agrees with this comment, and a discussion of the relative importance of the sources of uncertainty will be provided in the revised ERA.

Ralph G. Stahl:

The significant uncertainties in the analysis of the small mammal assessment endpoint were identified and adequately addressed. I do not believe the significant uncertainty in the assessment of the red fox was clearly identified nor adequately addressed in the ERA. The assessment for the red fox could be improved by reviewing data on dogs or other canines exposed to tPCBs or TEQ. This was not done nor the possibility of needing to do so discussed in the ERA.

RESPONSE 3.7-RS-6

EPA was unable to locate any canine toxicity data for tPCBs or TEQ in the review of the effects literature performed while developing the effects data set for the ERA. EPA will update the literature review to determine if recent canine studies have been published, and any such studies will be included in the revised ERA.

Timothy Thompson:

Within the limits of the data, uncertainties fairly well characterized. Same general previous comments re. uncertainty apply here as well.

3.7(h) Was the weight of evidence analysis appropriate under the evaluation criteria? If not, how could it be improved?

Valery Forbes:

The WOE appears appropriate, but see comment under 3.4.h on ambiguity of WOE scoring sheets. In addition, there should be separate WOE tables for shrews and red fox due to the different kinds of information available for each (see e.g., Table 4.7).

RESPONSE 3.7-VF-2

Please refer to the responses to General Issues 8 and 16.A.

Thomas W. La Point:

This is a particular instance, in this reviewer's opinion, where the WOE approach does not work. When all avenues of evidence are highly "suspect" because of limited quantification, limited sampling success, and high uncertainty, the WOE approach should have an option to "go for more data." As it is, given the large number of mouse and shrew species, given the negative regression of shrew mortality and soil PCB concentration, given the well-published results indicating reproductive effects stemming from chronic PCB exposure in small mammals, I would recommend not putting much weight on the WOE approach and employ the results of the Monte Carlo and probability bounds modeling and the limited – but insightful – results of the GE field study.

To lessen risk estimates because of so much uncertainty is neither wise nor prudent. This situation calls for either more data or more conservatism.

RESPONSE 3.7-TL-4

EPA believes that the WOE approach is the appropriate process for combining different types of information on risk in an objective and unbiased fashion. For short-tailed shrews, the WOE approach did rely on the results of the Monte Carlo and probability bounds modeling and the results of the GE field study, as requested by the Reviewer. However, EPA also included the results of the field surveys in the WOE analysis because this study provided important information on the approximate abundance and reproductive history of short-tailed shrews and other small mammals in the PSA. Collection of additional data is not possible because of practical limitations.

James T. Oris:

See previous comments on WOE analysis.

Mary Ann Ottinger:

There are inconsistencies in the field assessments and in the WOE, which need clarification.

RESPONSE 3.7-MO-6

Please refer to the response to General Issue 8.

Brad Sample:

The WoE was conducted for omnivorous/carnivorous mammals as a whole and not separately for shrews and fox. Because the same data were not available for each receptor and because the shrew and fox represent distinct exposure pathways, separate WoE analyses should be performed for each receptor.

RESPONSE 3.7-BS-13

EPA agrees; please refer to the response to General Issue 16.A.

Ralph G. Stahl:

The weight of evidence analysis for small mammals was appropriate under the evaluation criteria. Because of the problems associated with the exposure and effects assessment for the red fox, the weight of evidence analysis for the red fox should be separated from that of the shrew, and, even better, deleted from the ERA.

RESPONSE 3.7-RS-7

Please refer to Response 3.7-RS-5 and the response to General Issue 16.A.

Timothy Thompson:

Yes – although I would recommend that a separate table be constructed for foxes and for shrews.

Response 3.7-TT-3

EPA agrees; please refer to the response to General Issue 16.A.

3.7(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?

Valery Forbes:

No estimates of risk were made outside of the PSA.

Thomas W. La Point:

Yes, but the risk estimates are so broad as to not be very useful. The limited data (literature and field study re-analysis) would indicate a greater risk than indicated by the WOE approach.

Given the large number (40+) of small mammal species on the PSA, conservatism at Locations 13 and 14 would seem to be called for.

RESPONSE 3.7-TL-5

In the revised ERA, risk conclusions will be specific to the representative species and, for short-tailed shrews, will be evaluated in the context of contaminant concentrations throughout the PSA, not just at the locations sampled.

James T. Oris:

Yes.

Mary Ann Ottinger:

The lower reaches should pose reduced threat due to diminishing contamination.

Brad Sample:

A downstream evaluation not done for omnivorous/carnivorous mammals. Rather extrapolation to other omnivorous/carnivorous mammal species was provided instead. Some explanation should be provided to explain why downstream risks were not evaluated.

RESPONSE 3.7-BS-14

A downstream assessment was not conducted for omnivorous and carnivorous mammals primarily because a site-specific MATC was not calculated, and PCB concentrations in the floodplain decline significantly below the PSA. Further explanation of the potential for downstream risks will be provided in the revised ERA.

Ralph G. Stahl:

The risk estimates were appropriately derived for reaches of the river where site-specific studies were not conducted.

Timothy Thompson:

No estimates of risk were made outside of the PSA. If those are developed for the final ERA, the same general rules for MATC development apply here.

3.7(j) In the Panel members' opinions, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?

Valery Forbes:

The ERA concluded that risks to omnivorous and carnivorous mammals is moderate to high but that there is uncertainty in these conclusions due lack of data. I agree with this conclusion.

Thomas W. La Point:

No. See my responses to 3.7(g), 3.7(h), and 3.7(i) for justification of my opinion.

James T. Oris:

Yes, supported.

Mary Ann Ottinger:

The endpoints measured and some flaws in the design of the field studies limit data interpretation. Risk assessment follows from the available information.

Brad Sample:

Although some clarification and revision of the analyses performed are needed, I believe that the conclusions are probably appropriate and are supported by the available data.

Additional discussion is also needed to indicate that the evaluation is for the local population and not for individuals.

RESPONSE 3.7-BS-15

The conclusions regarding risks of tPCBs and TEQ to omnivorous and carnivorous mammals apply to the locally exposed population. Please refer also to the response to General Issue 9.

Ralph G. Stahl:

The evaluation supports the conclusions regarding intermediate risk to small mammals, although pending the outcome of a singular, mutually agreeable statistical approach, there could be some modification to this designation.

The evaluation does not support the conclusions regarding intermediate risk to red fox. The evaluation of effects was not appropriate under the evaluation criteria and, in my opinion, was not scientifically defensible. The conclusions are insupportable currently and I recommend that the red fox assessment endpoint be deleted from the ERA.

RESPONSE 3.7-RS-8

Please refer to Response 3.7-RS-5 above.

Timothy Thompson:

Yes. Given the site-specific data, along with an abundance of literature data, sufficient information exists upon which to develop and MATC and formulate an appropriate risk management decision for this receptor group.

3.8 Threatened and Endangered Species

> ***3.8.(a) Were the EPA studies and analyses performed (e.g., field studies, site-specific toxicity studies, comparison of exposure and effects) appropriate under the evaluation criteria, and based on accepted scientific practices?***

Valery Forbes:

Yes as far as I can tell.

Thomas W. La Point:

Yes. There was an emphasis on modeling, as no bald eagles or bitterns were observed during the surveys, despite having observations of their presence seasonally.

RESPONSE 3.8-TL-1

Bald eagles and American bitterns were both sighted repeatedly in the PSA over a several-year period. Observation records are detailed in Appendix A.1 to the ERA. Bitterns are believed to attempt to breed in the PSA, although eagles currently do not. Eagles attempted to breed in the PSA in the 1990s and they also nest downstream on the river in CT. Modeling was used for bitterns and eagles because they are threatened and endangered species, which severely limits opportunities for collection of tissue samples to estimate body burdens. Also, both of these species have large home ranges and are territorial, which translates into small sample sizes when they are present. For these reasons, modeling is the best tool to estimate exposure and effects for eagles and bitterns.

James T. Oris:

Within the limits of the study designs, the studies were appropriate.

Mary Ann Ottinger:

Representative species were the bald eagle, American bittern, and small-footed myotis; eagles and bitterns have been observed in the PSA. Models were based on prey item tPCBs and estimated sensitivities, based on the literature from related species.

Brad Sample:

Studies and analyses performed by EPA for T&E species were generally suitable and appropriate.

Ralph G. Stahl:

There were no spite-specific toxicity studies and very limited field studies conducted by EPA. The limited field studies (playback calls, anabat study, eagle surveys) were appropriate under the evaluation criteria and based on accepted scientific practices. As noted in the ERA however, these field evaluations were not designed to be quantitative and therefore were not included in the weight-of-evidence evaluation of risk.

The comparison of exposure and effects was based entirely on modeled information which inserts a high degree of uncertainty in the analysis.

The use of stomach contents from the tree swallow study to estimate tPCB exposure to the small footed myotis presents a high degree of uncertainty in this portion of the analysis. Estimates of exposure for the bald eagle are based on prey (fish, waterfowl, small mammals) data collected from the PSA. Estimates of exposure for the American bittern are based on prey (amphibians, insects, crawfish) data collected from the PSA. In this regard, the estimates of exposure for the bald eagle and the American bittern are appropriate under the evaluation criteria.

RESPONSE 3.8-RS-1

Tree swallows and small-footed myotis are both aerial insectivores that capture flying insects over water. Both of these species were observed feeding on emerging aquatic insects in the same areas of the river; however, the swallows were seen feeding during the day, whereas the bats were seen feeding at night. There is a high likelihood that the emerging insects that both species eat have similar tPCB body burdens, even though some of them are different species. EPA recognized that the use of stomach contents from the tree swallow study to estimate exposure to small-footed myotis has some uncertainty, as described in Appendix K to the ERA

Timothy Thompson:

Yes – modeling issues have been dealt with previously

3.8(b) Were the GE studies and analyses performed outside of the framework of the ERA and EPA review (e.g., field studies) appropriate under the evaluation criteria, based on accepted scientific practices, and incorporated, appropriately in the ERA?

Valery Forbes:

None performed.

Thomas W. La Point:

N.A. The field study conducted was a survey to look for the presence of eagles, bitterns and s.f. myotis (among other T&E species).

James T. Oris:

None were performed

Mary Ann Ottinger:

Surveys would consider these species, but data were not available.

Brad Sample:

No studies were performed by GE for T&E species.

Ralph G. Stahl:

There were no GE studies conducted for this assessment endpoint.

Timothy Thompson:

No GE studies were performed for this receptor group.

3.8(c) Were the estimates of exposure appropriate under the evaluation criteria, and was the refinement of analyses for the contaminants of concern (COCs) for each assessment appropriate?

Valery Forbes:

Yes they seem so.

Thomas W. La Point:

The use of literature data and the concentrations in potential prey items was appropriate.

James T. Oris:

EPA should consider how habitat use may affect the exposure assessment outcomes. Is it realistic to assume utilization by eagles of only Woods Pond?

RESPONSE 3.8-JO-1

Suitable bald eagle habitat exists in the southern sections of the PSA below New Lenox Road (Reaches 5C, 5D, and 6), particularly the large open-water areas of Woods Pond and its associated backwaters. Bald eagles were observed in the PSA foraging over these water bodies and roosting in large trees along the shoreline. The assumed exposure area for bald eagles was the southern portion of the PSA, from the downstream portion of Reach 5B to Woods Pond. This area was not subdivided for the exposure assessment because bald eagles would likely forage throughout this area (see Section 11.2 of the ERA).

Can the diet of bats realistically be extrapolated from the swallow diet, even though they occupy temporally different niches?

RESPONSE 3.8-JO-2

Bats and swallows are aerial insectivores that have been observed feeding in the same habitats and spatial niches in the PSA, although at different time periods; bats feeding from dusk through the early hours of the morning and swallows feeding primarily between sunrise and sunset. Little is known about the specific feeding habits of the small-footed myotis. Small-footed myotis were assumed to have feeding habits similar to other bats observed in the PSA, which were feeding on insects over and along the riparian zones of the river. These are the same areas used by swallows. Although there is some uncertainty associated with this assumption, EPA believes it is defensible to extrapolate exposure to myotis from the swallow diet.

Mary Ann Ottinger:

COCs were estimates based on prey tPCBs and species sensitivity estimates. This was appropriate given the limitation of access to these animals.

Brad Sample:

Exposure modeling for the T&E species was conducted in the same way as for other wildlife receptors. Consequently many of the issues previously discussed for other wildlife receptors (i.e., reliance on measured versus modeled tissue concentrations; degree to which measured data represent full range of potential exposure in the PSA, etc.) also apply for the T&E species.

RESPONSE 3.8-BS-1

For bald eagles and American bitterns, the prey data represent the full range of potential exposure in the PSA. For small-footed myotis, EPA agrees with the Reviewer and will expand the analysis to include the full range of potential exposure concentrations in the PSA.

Other specific issues include:

- Exposure estimation for Bald Eagles - The fish component of the eagle diet is broken down by feeding guilds (predatory, forage, and bottom feeders). Although this is correct, I believe the size-based approach outlined above for other piscivorous wildlife will provide a more suitable representation of exposure for bald eagles.

RESPONSE 3.8-BS-2

EPA will conduct a sensitivity analysis for the prey size-based approach and will include the results in the uncertainty section of the revised ERA.

- Use of the egg accumulation model for bald eagles is appropriate given the T&E status of the species. However because uncertainty associated with the model is extremely high, some effort to validate the modeling approach should be performed. This would mean identifying at least one location (preferably more than one) anywhere within the range of the bald eagle for which there exist data on measured concentrations in bald eagle eggs plus associated fish concentration data that can be used to estimate dietary exposure. To allow for the correct interpretation of this modeling approach, some information on the accuracy of the resulting estimate is needed.

RESPONSE 3.8-BS-3

EPA will investigate the availability of data that could be used to conduct this evaluation, and give consideration to validating the modeled data. This information, if available, will be incorporated into the revised ERA.

- In addition, it needs to be noted that bald eagle do not currently nest in the PSA.

RESPONSE 3.8-BS-4

As described in Appendix K, Section 2.1.5, bald eagles were chosen as one of the representative T&E species because:

- The bald eagle is the only federally listed species that was observed in the PSA.
- Habitat within the PSA is suitable for bald eagles.
- They have attempted to breed within the PSA and could do so in the future.
- The prey they consume are directly exposed to COCs in the area.

Suitable bald eagle habitat is found in the southern sections of the PSA below New Lenox Road (Reaches 5C, 5D, and 6), particularly the large open-water areas of Woods Pond and its associated backwaters. Bald eagles were observed in the PSA foraging over these water bodies and roosting in large trees along the shoreline. A bald eagle pair attempted to nest in the PSA near Woods Pond in the mid-1990s; however, the nest was destroyed in an April snowstorm and the pair did not attempt to re-nest that year or subsequently.

- This analyses simply evaluates potential effects that could occur should individuals start to nest at some point in the future.

- Note – the egg exposure modeling approach used for bald eagles could also have been applied to the other avian receptors that nest in the PSA (validation for these species would also be needed to calibrate he accuracy of the method). This would have added an additional measure to evaluate exposure and effects.

RESPONSE 3.8-BS-5

EPA agrees with this comment and is considering using the egg exposure modeling approach for other avian species in the revised ERA.

- The diet of the bat was assumed to be adequately represented by the stomach content data from nestling swallows. Although this may be correct, additional analyses and discussion are needed to support this assumption. In particular, comparisons relating diet for swallow (unpubl. Custer data) to that of the bat (literature) should be performed.

RESPONSE 3.8-BS-6

EPA believes that the tree swallow stomach content data are applicable for estimating dietary exposure for small-footed myotis. EPA will provide additional descriptions of bat and swallow diets in the revised ERA.

- References for sources of specific exposure model parameter values need to be provided in Tables K.2-2 and K.2-3, Tables K.2-10 and K.2-11, and in Tables K.2-18 and K.2-19.

RESPONSE 3.8-BS-7

EPA agrees, and will provide the requested references in the revised ERA.

Ralph G. Stahl:

Estimates of exposure were heavily based on modeling although site-specific prey item data were utilized in the assessment.

The use of stomach contents from the tree swallow to represent potential exposure to tPCBs from prey insects consumed by the small footed myotis inserts a high degree of uncertainty into the exposure estimates. The tPCB analysis of the insects represents the only site-specific exposure information available to make this estimate, yet it is uncertain whether small footed myotis consume the same insects as found in tree swallows. The exposure assessment for the small footed myotis was appropriate under the evaluation criteria.

RESPONSE 3.8-RS-2

Please refer to Response 3.8-BS-6.

The use of site-specific fish, waterfowl and small mammal tissue tPCB and TEQ content was appropriate for the estimate of exposure, via diet, to the bald eagle. The statistical treatment of these exposure data is complex and not all together well understood. After re-reading both the main text of the ERA and Appendix K there is still a high degree of confusion as to the strengths and weaknesses of the statistical treatment (Land H-statistic). I recommend that the authors of the ERA revisit the statistical discussion to determine if there are ways that the description can be simplified. The exposure assessment for the bald eagle was appropriate under the evaluation criteria.

RESPONSE 3.8-RS-3

The presentation and discussion of statistical methods will be expanded and/or simplified as appropriate throughout the revised ERA. Please refer to the response to General Issue 1.

The use of site-specific tPCB, TEQ from amphibians, benthic insects and crawfish was appropriate for estimating exposure in prey items for the American bittern. The exposure assessment for the American bittern was appropriate under the evaluation criteria.

Timothy Thompson:

Yes.

3.8(d) Were the effects metrics that were identified and used appropriate under the evaluation criteria?

Valery Forbes:

There seem to be very important uncertainties here, given the lack of relevant effects data for T&E species.

RESPONSE 3.8-VF-1

Much of the uncertainty related to the T&E species chosen for inclusion in the ERA is associated with the lack of species-specific toxicology data. Because threatened and endangered species are in danger of extirpation and typically have small numbers of individuals available for study, there are fewer effects data in the literature than for other species. Species-specific effects data for bald eagles were used in the risk assessment when estimating effects to eggs. Data for great blue herons and black-crowned night-herons, however, were used to estimate risks to American bitterns. Toxicity threshold ranges were developed for both bald eagles and American bittern. No toxicology data were available for the small-footed myotis. Although there were several studies done on little brown bats, there were insufficient data available to develop a threshold effects range. Data were available, however, to derive dose-response curves using surrogate mammals for the small-footed myotis. These uncertainties were described in the ERA.

Thomas W. La Point:

Yes.

James T. Oris:

Effects metrics were defined appropriately

Mary Ann Ottinger:

There were few data collected at the PSA. Maternal deposition of tPCBs into the egg was used as an estimate of embryonic exposure. This is appropriate due to known sensitivity of raptors to PCBs.

Brad Sample:

Recognizing the limitations in available toxicity data, the effects metrics selected are suitable and appropriate.

Ralph G. Stahl:

The effects metrics generally were appropriate under the evaluation criteria. However, the primary basis for determining effects to bald eagle and American bitterns was based on literature-based studies in domestic species (chickens, pheasants), kestrels, herons or similar species. The relative sensitivity of bald eagles and American bittern compared to the species tested in toxicological evaluations is not known, but is likely to be lower than those evaluated in these laboratory studies. The effects metrics for the bald eagle and the American bittern were appropriate under the evaluation criteria.

In the case of the small footed myotis, toxicological evaluations were available for the big brown bat and little brown bat. These species are likely to be relatively close in physiology, metabolism and perhaps sensitivity to organochlorine chemicals as the small footed myotis. The effects metrics for the small footed myotis were appropriate under the evaluation criteria.

Timothy Thompson:

Within the limits of modeling discussed in previous section, the evaluation criteria are appropriate.

3.8(e) Were the statistical techniques used clearly described, appropriate, and properly applied for the objectives of the analysis?

Valery Forbes:

No additional comments.

Thomas W. La Point:

Yes. This was a good use of probabilistic modeling.

James T. Oris:

See comments in 3.4

Mary Ann Ottinger:

The models are based on available information and similar species, where possible. It is appropriate to use the range of sensitivities to tPCBs determined in lab and field studies for the models. This increases the likelihood that the true sensitivity of an organism is within the range of sensitivities and is warranted because the loss of even one individual of an endangered species is critical to the population.

Brad Sample:

The issues with the probabilistic analyses previously mentioned under other receptors apply for the analyses for T&E species.

Ralph G. Stahl:

The statistical techniques were clearly described, appropriate and properly applied for the objectives of the analysis.

Timothy Thompson:

Within the limits of modeling discussed in previous section, the evaluation criteria are appropriate.

3.8(f) Was the characterization of risk supported by the available information, and was the characterization appropriate under the evaluation criteria?

Valery Forbes:

If it can be assumed that T&E species are such, not because they are particularly sensitive to PCBs and other toxicants, but rather because they have life-history characteristics that make them sensitive to demographic impacts, then it is appropriate to use effects thresholds from non T&E species in the assessment as was done here. However it should be recognized that the same impairment of survival or reproduction could have much greater consequences for T&E populations than for non T&E populations.

RESPONSE 3.8-VF-2

EPA agrees that impairment of survival or reproduction could have much greater consequences for T&E populations, and thus established more conservative rules for interpreting the risks for these species.

There was a problem with the risk estimates differing for different life stages of the bald eagle and an average taken as an estimate of overall risk. This is inappropriate as discussed under Charge Question 2. A life cycle model could have been applied to calculate elasticities (sensu Caswell) of the different life stages to quantify their relative importance to population dynamics of this species.

RESPONSE 3.8-VF-3

In the revised ERA, risks to bald eagles will be based on estimated effects to bald eagle eggs.

Thomas W. La Point:

Yes.

James T. Oris:

There is some confusion in how the risk was characterized for eagles. Considering that the risk was determined to be “high” for eggs and “low” for adults, how is it possible that the overall risk is “moderate”? It would seem appropriate that the risk should be determined based on the most sensitive life stage. Loss of fish-eating bird populations due to DDT exposure was because of impacts in eggs and not directly on adults, but the risks were high and the impacts were severe. It makes no sense to take the apparent average of the two risk characterizations between eggs and adult.

RESPONSE 3.8-JO-3

In the revised ERA, risks to bald eagles will be based on estimated effects to bald eagle eggs.

Mary Ann Ottinger:

The risk characterization is supported by the literature to the extent possible.

Brad Sample:

The risk characterization is suitable and appropriate.

Organizational note: descriptions of the field studies and their analyses only appear in the Risk Characterization section. As these are really effects data, they should be described in the Effects section, with interpretation and integration of results left in the Risk Characterization section. This anomalous presentation of data also occurs for some other receptors. The organization of the report would benefit if the presentation of the data were moved.

RESPONSE 3.8-BS-8

EPA will review the presentation of descriptions of field studies and their analyses and make the recommended organizational changes to the Risk Characterization section in the revised ERA.

Ralph G. Stahl:

The characterization of risk was not fully supported by the available information, and carries a high degree of uncertainty due to the characterization relying entirely on modeling. American bittern are known to inhabit the PSA according to the ERA, and bald eagles are known to nest in the Housatonic River watershed downstream of the PSA in Connecticut. It is also known that the bald eagle has attempted to nest in the PSA; however, whether this was successful or failed is not clearly articulated in the ERA. Further, if the nesting attempts failed, there is no discussion as to whether this resulted from exposure to tPCBs or TEQ in the watershed.

RESPONSE 3.8-RS-4

As described in Appendix K, Section 2.1.5, a bald eagle pair attempted to nest in the PSA near Woods Pond in the mid-1990s; however, the nest was destroyed in an April snowstorm and the pair did not attempt to re-nest that year or subsequently.

There is evidence that 2 of the 3 T&E species under consideration are found in the watershed and appear to be either utilizing the area and/or nesting there. In contrast to the modeled exposure and effects evaluations, the limited survey data appear to indicate that these two species are not significantly impacted by the tPCBs or TEQ.

RESPONSE 3.8-RS-5

The evidence/observation of individual threatened or endangered species in an area does not infer absence of risk, particularly if the species is migratory and has a large home range, as is the case for the bald eagle and American bittern, and successful reproduction within the PSA has not been observed. Where there is limited observational data, presence or absence information and timing of the observations might be used to infer risk only if an acute lethal effect is associated with a COC. All three T&E species evaluated in the ERA are migratory, and it is likely that individuals are moving into the PSA from uncontaminated areas. Barring an acute lethal effect, these same individuals could establish residence in the PSA and consume contaminated prey, as was modeled, and be negatively affected during the reproductive period. EPA believes that the limited T&E survey data are not sufficient to draw a conclusion that there are no effects from COCs in the PSA.

Timothy Thompson:

Yes, within the limits of uncertainty the assessment of risk is appropriate.

3.8(g) Were the significant uncertainties in the analysis of the assessment endpoints identified and adequately addressed? If not, summarize what improvements could be made.

Valery Forbes:

Yes, and although they are large, I cannot see how they can be reduced further without collecting additional data.

Thomas W. La Point:

Yes. The large uncertainties arise from the lack of specific PSA data (presence, feeding, territory size and use) on any of the T&E species.

James T. Oris:

Uncertainties were appropriately described, but more discussion is needed in terms of analytical uncertainties

RESPONSE 3.8-JO-4

EPA evaluated the variability in analytical data and presented the results in ERA Appendix C.11. This information will be presented in more detail in the discussion of uncertainty in the revised ERA.

Mary Ann Ottinger:

Uncertainties are associated with the model development due to the use of information developed from other species. As most of this information is from peer reviewed journal papers, there is some level of confidence in the rigor of those data.

Brad Sample:

Although the uncertainties are generally well captured in the ERA report, the relative importance of various uncertainties is not discussed. Interpretation and understanding of the results would benefit by such an analysis. Spatial heterogeneity of contamination and the degree to which the exposure data actually represent the PSA are very significant issues. These need to be discussed.

RESPONSE 3.8-BS-9

EPA agrees with the comment regarding the relative importance of the sources of uncertainty, and a discussion of the relative uncertainties will be provided in the revised ERA. EPA will also provide a discussion of the representativeness of the data used in the exposure analyses relative to the PSA in the revised ERA.

Ralph G. Stahl:

The uncertainties in the analysis of the assessment endpoints were clearly identified and adequately addressed. As noted earlier the fact that the entire exposure and effects assessment were based on modeling inserts a high degree of uncertainty into the analysis.

Timothy Thompson:

The final ERA should address how the limits of analytical variability could affect the overall determination of risk.

RESPONSE 3.8-TT-1

EPA agrees and will discuss this issue in the Sources of Uncertainty section for all Assessment Endpoints.

3.8(h) Was the weight of evidence analysis appropriate under the evaluation criteria? If not, how could it be improved?

Valery Forbes:

It is extremely limited due to lack of data.

Thomas W. La Point:

Yes. One way to improve the WOE analysis would be to band several eagles and/or other T&E species and follow them for several years. Another way to improve the WOE would be to use surrogate species (mergansers) to study PCB and TEQ uptake from the fish in the PSA. However, please note my comment in 8.j.

RESPONSE 3.8-TL-2

While there are undoubtedly additional studies that could be performed to gather supplemental information for use in the risk assessment, EPA believes that the estimated exposure for bald eagles was robust and appropriate because it was based on site-specific exposure data and a number of peer-reviewed studies for eagles and tPCBs/TEQ, as described in Appendix K, Section 2.1.5.1.

James T. Oris:

See previous comments on WOE analysis.

Mary Ann Ottinger:

The WOE depends on models generated from the literature and are predictive given the data from the content of tPCBs in their diet.

Brad Sample:

The WoE for T&E species was conducted as if each T&E species was an independent line of evidence. This is inappropriate. Because the conceptual models and exposure represented by each species differ, a separate WoE evaluation should be performed for each species.

RESPONSE 3.8-BS-10

As described in Section 11.4.3 of the ERA, a separate WOE evaluation was performed for each species, although data for these evaluations were placed in the same tables. The WOE evaluations will be presented separately in the revised ERA.

As stated previously, the weight-of-evidence (WoE) analyses is not adequately transparent. The WoE analyses use best professional judgment to apply final weightings for the various lines of evidence and their attributes. It is unclear how final weightings were derived. For example, it is not clear why evidence for harm for small footed myotis from PCBs and TEQs and for the bittern for TEQs are listed as undetermined. Some discussion as to how these weights were derived needs to be provided – were all attributes judged equally? If not need to know why and this needs to be explained – this will allow reviewers to judge the process.

RESPONSE 3.8-BS-11

Please refer to the response to General Issue 8.

Ralph G. Stahl:

The weight of evidence analysis was not fully appropriate under the evaluation criteria. One concern is the fact that the site-specific biological surveys were excluded from this analysis. This does not appear to be objective, reasonable or consistent with evaluations for other receptors.

The reason given in Appendix K for not including this survey information is that the studies were not quantitative in nature; however, given the significance of the risk question at hand it is important that all available information be considered to the maximum extent possible. These are the only site-specific data which can be useful in evaluating the potential for risk. The presence of the T&E species I the PSA and the Housatonic River watershed is crucial to understanding whether the tPCBs and TEQ exposure may be problematic. If the T&E species are found to inhabit the PSA, or use it for feeding, resting, etc. then there is some evidence that the exposures to tPCBs and TEQ are not sufficiently problematic as to impair the utilization of the area by these important species.

RESPONSE 3.8-RS-6

The evidence/observation of individual threatened or endangered species in an area does not infer absence of risk, particularly if the species is migratory and has a large home range, as is the case for the bald eagle and American bittern, and successful reproduction within the PSA has not been observed. Where there is limited observational data, presence or absence information and timing of the observations might be used to infer risk only if an acute lethal effect is associated with a COC. All three T&E species evaluated in the ERA are migratory, and it is likely that individuals are moving into the PSA from uncontaminated areas. Barring an acute lethal effect, these same individuals could establish residence in the PSA and consume contaminated prey, as was modeled, and be negatively affected during the reproductive period. EPA believes that the limited T&E survey data are not sufficient to draw a conclusion that there are no effects from COCs in the PSA.

Timothy Thompson:

Yes. However, I concur with my colleagues who noted that there should be different risk weighting for different life stages. Particularly in the fact that eggs were at high risk, adults at low risk, and the final assignment was "moderate". This is problematic in that eggs are a more sensitive life stage than adults.

RESPONSE 3.8-TT-2

EPA will expand the text in the revised ERA to clarify that the overall risk characterization for eagles is based on an estimate of high risk to eggs.

3.8(i) Were the risk estimates objectively and appropriately derived for reaches of the river where site-specific studies were not conducted?

Valery Forbes:

Appears ok.

Thomas W. La Point:

Somewhat. If the osprey is considered at high risk, but uncertain because only one line of evidence was employed, why not the same conclusion for the bald eagle and bittern? Ospreys and bald eagles are competitors in habitats in which they coexist. Hence, a logically-applied WOE would conclude very similar risk levels for both species.

RESPONSE 3.8-TL-3

In response to the Reviewers' comments, the WOE assessments will be revised for bald eagles, American bitterns, and other species to make them concordant and more transparent. EPA will also describe the sources and the magnitude of uncertainty in greater detail in the revised ERA. Please refer also to the response to General Issue 8.

James T. Oris:

Yes.

Mary Ann Ottinger:

Yes, based on the data available.

Brad Sample:

A downstream evaluation not done for T&E species. Rather extrapolation to other species taxonomically similar to the T&E species was provided instead. Some explanation should be provided to explain why downstream risks were not evaluated.

RESPONSE 3.8-BS-12

A downstream evaluation was performed for the bald eagle. Appendix K, Section 4.2.1.5, contains a detailed description of risk estimates downstream of Woods Pond. Downstream risks were evaluated for bald eagles by evaluating concentrations of tPCBs in prey items in Reaches 7 to 16 and then comparing these concentrations to the MATC developed for bald eagles.

Ralph G. Stahl:

The risk estimates were appropriate for reaches of the river where site-specific studies were not conducted.

Timothy Thompson:

Yes. My only comment is that the fact that suitable habitat exists for eagle nesting, but nests are not found within the PSA is not in and of itself an indicator of risk. This can, however, be used as a pillar supporting the overall assessment of risk.

RESPONSE 3.8-TT-3

The absence of an active bald eagle nest in the PSA was not considered in the assessment of risk. Risk was estimated for eagles that are likely to nest in the PSA in the near future. Bald eagle populations in the northeast have been expanding over the past 20 years as they recover from the effects of DDE exposure. During the expansion, eagles first reoccupied former habitat along the coast, and have since occupied the larger river systems. This dispersal effect has been observed along the Housatonic River, and eagles currently nest along the river in CT.

As stated in Appendix K, Section 2.1.5, eagles were chosen as one of the representative T&E species because the bald eagle is the only federally listed species that was observed in the PSA, habitat within the PSA is suitable for bald eagles, they have attempted to breed within the PSA and could do so in the future, and the prey they consume are directly exposed to COCs in the area. There is a high likelihood that eagles will attempt to nest in the PSA in the future. A bald eagle pair attempted to nest in the PSA near Woods Pond in the mid-1990s; however, the nest was destroyed in an April snowstorm and the pair did not attempt to re-nest that year or subsequently.

3.8(j) In the Panel members' opinions, based upon the information provided in the ERA, does the evaluation support the conclusions regarding risk to local populations of ecological receptors?

Valery Forbes:

The ERA concludes that the risk is high for selected threatened and endangered species but that there is uncertainty in this conclusion. I would rate the risk as very uncertain given the heavy emphasis on modelling results (which for all wildlife gave higher risk estimates than field or site-specific data where these were available).

RESPONSE 3.8-VF-4

EPA will describe the sources and the magnitude of uncertainty in greater detail in the revised ERA.

Thomas W. La Point:

Yes. I note that in this instance, with large charismatic T&E species, that the WOE approach fully uses and relies on modeling. If the same approach were to be used for piscivorous and omnivorous mammals, the reliance on limited field data would be lessened and greater weight put on the modeling. If the same "weighting" on high-variance, small sample field testing were to be applied here, it would moderate the "moderately high" risk determination by lessening it to "moderate." Could it be that shrews and foxes are not he ld in as high esteem as bald eagles?

RESPONSE 3.8-TL-4

The weighting of any line of evidence was done independently and was not affected by the numbers of lines of evidence available or the species being assessed. The comparison of modeled exposure to the effects line of evidence received different weights between wildlife species because the quantity and quality of data available to estimate risk varied from species to species.

James T. Oris:

Given the limitations in the studies and the high uncertainties associated with the extrapolations used, the conclusions are supported. However, the EPA needs to clearly articulate the high uncertainty here.

RESPONSE 3.8-JO-5

EPA acknowledged the high uncertainty associated with this Assessment Endpoint in the ERA (Section K.4.4).

Mary Ann Ottinger:

Risk assessment appears to be reasonable given available information

Brad Sample:

Yes, I believe that the conclusions are generally supported by the analyses.

Ralph G. Stahl:

The information provided in the ERA generally supports the conclusions regarding moderate, but not high risk to T&E species.

RESPONSE 3.8-RS-7

In response to the Reviewers' comments, EPA will revisit the weight-of-evidence evaluation and will separate the risk estimates for T&E species in the revised ERA.

Timothy Thompson:

Within the limits of modeling uncertainty, yes.

4. SUMMARY DISCUSSIONS AND CONCLUSIONS IN THE ERA

4. Are the summary discussions and conclusions in the ERA supported by the information provided in the report, and did the conclusions describe the risks in an objective, reasonable, and appropriate manner?

Valery Forbes:

General Comment on WOE - In their ecological risk assessment guidelines, EPA (1998) states: "A major advantage of field surveys is that they provide a reality check on other risk estimates, since field surveys are usually more representative of both exposures and effects (including secondary effects) found in natural systems than are estimates generated from laboratory studies or theoretical models." Thus it should not be particularly surprising that field surveys do not always match the risks predicted from laboratory studies or models. The above statement would support weighting field studies higher than either lab or model outputs, when these provide conflicting estimates of risk. Although field studies also have their limitations these need to be carefully judged in comparison with the limitations of lab/mode l extrapolations. I am not sure that this has been consistently and objectively done in the present ERA. Given the subjectivity in the qualitative WOE used here, I recommend that the relative weightings of laboratory, literature, model, and field results be given further review.

RESPONSE 4-VF-1

EPA (1997) recognizes that there can be limitations associated with field studies of populations and/or communities. In addition, the guidance suggests relying on multiple lines of evidence, and does not recommend that one line take precedence over another. Therefore, EPA believes that no single line of evidence should be given more weight a priori, and that each measurement endpoint should be evaluated separately and on its own merits, as was done in the ERA.

However, based upon the recommendation of the Reviewers (who are not always in agreement with the weight that should be placed upon field studies), EPA will reevaluate the weighting of each measurement endpoint, particularly field studies, in the revised ERA. Please refer to the response to General Issue 8.

The HQ analyses should be modified to consider the size of the HQ (not simply is it greater than one?) and some measure of probability or proportion of samples exceeding the critical HQ of 1.

RESPONSE 4-VF-2

The use and interpretation of HQs in the ERA is discussed in the response to General Issue 7.

In my view the issue of risk acceptability should not be addressed in the ERA. The ERA should focus on quantifying risks and their associated uncertainties. Acceptability is a risk management issue.

RESPONSE 4-VF-3

EPA recognizes the Reviewer's comment and will remove the discussion of whether risks are acceptable or unacceptable in the revised ERA, although EPA (1998) suggests that there is a need for such information to be included in the ERA. For further discussion of this topic, please refer to the response to General Issue 7.

I would recommend that use of the term 'significant' (e.g., as in significant risk) be restricted to situations in which reference is being made to statistical tests.

RESPONSE 4-VF-4

EPA agrees with this comment; please see General Issue 7.

Thomas W. La Point:

Mostly. There are inconsistencies in approaches for the piscivorous birds and the T&E representative, the bald eagle. As an idea and in my opinion, the bald eagle (or mink) results could be used as a surrogate measure for the ospreys, kingfishers, and other predominately fish-eating vertebrate fauna. The ERA acknowledges that there are substantial differences in species sensitivities to COCs (particularly the PCBs). However, given the value of using surrogate species, species chosen to represent other species present in the PSA, care should be taken to be conservative, as it is most certain that several un-measured species are more sensitive than are the ones selected to represent the biotic community (e.g., data presented in Table 12.4-1).

RESPONSE 4-TL-1

EPA concurs with the Reviewer's comment there may be more sensitive species than the representative species selected for the Assessment Endpoints.

In the Housatonic River ERA, EPA's objective was to conduct an unbiased assessment of risks to biota. For example, rather than rely on only the most sensitive species in developing effects metrics when species-specific data were lacking, EPA developed threshold ranges that spanned effects thresholds from the most sensitive to the most tolerant species studied.

In the Conclusion Section 12.2.1.3 ("Fish"), the statement concludes, "The magnitude of impact is not predicted to be catastrophic in any reach;" I would propose that the manner in which PCBs exert their toxicity is not in a "catastrophic" manner, but in chronic effects on reproductive behavior. Further such reproductive effects will ultimately influence the survivorship of predators feeding on fish in the PSA – hence, conservatism should be called for. Subtle responses require more samples, longer study durations and truly focused studies on the critical questions ("would a full life-cycle test with trout or bass indicate reproductive effects in the F1 or F2 generations?").

RESPONSE 4-TL-2

The term "catastrophic" has been identified as a potentially value-laden term, and will be removed from the revised ERA. EPA used this term to indicate that the impacts to fish populations do not appear to be so large as to result in total reproductive failure of largemouth bass in the PSA, or to depend on migratory compensation to sustain a population. EPA agrees that alterations to reproductive behavior can result in such subtle responses. EPA also agrees that more focused studies would be required to definitively evaluate whether the observed reproductive effects in Housatonic River fish translate to more subtle population-level effects.

In almost every situation (insectivorous or carnivorous birds and mammals, for example), the field studies provided uncertainty in this ERA. This is not to say that those studies were poorly planned or conducted. They were good efforts to analyze population responses. However, as it is acknowledged that PCBs produce their affects over long-term exposure, the studies must have a duration and "intensity" to be able to quantify the subtle responses (a good example of a sufficiently-long study was the three-year swallow study). As evidenced by the complete reliance on modeling results for T&E species, the modeling results almost invariably result in estimates of greater risk. In my opinion, adding "clouding" and variance to these modeling estimates does not further the protectiveness of the risk assessment, particularly when the cost of conducting the focused studies is compared to the value of the species to be protected.

RESPONSE 4-TL-3

EPA agrees that additional research would contribute greatly to the understanding of risks from exposure to tPCBs and TEQ for various endpoints. However, given the practical limitations associated with the conduct of an ERA and guidance which states that it is not the goal of an ERA to embark upon research initiatives, these additional field studies will not be conducted, nor are they necessary to arrive at conclusions regarding risks to ecological receptors.

My comments above are particularly germane under the "Protection Goals" Section of the ERA (Section 12.4.2.1), in which broader aspects of the biotic community are to be protected as an ERA goal. Hence, the value of the mink, frog, osprey and eagle endpoints are not limited to protection of each species population, alone, but also as surrogates for the other species in the PSA. This section also calls for conservatism, as I would state we are not sure where "overly-conservative" estimates come in – but it is better to err closer to "over-conservative," than to err on the "upper end" of concentration limits.

RESPONSE 4-TL-4

EPA agrees that the species selected for each measurement endpoint are meant only to be representative of the other species that occur in the Housatonic watershed and that other species may or may not be at risk as well. The purpose of a discussion of relative risks to other species, which was provided in the risk characterization section for each endpoint, was to evaluate this issue, using known information about life history for the species. The Reviewers are not in agreement with regard to the degree of conservatism that should be used in the ERA.

As the ERA states (Pg. 12-70; line 31ff: "the central question….. is whether the exposed local populations are thriving in the contaminated habitat, not whether the larger regional population is surviving in spite of it." Because of the subtle, and decidedly un-catastrophic nature of the effects of PCBs, it must be kept in mind that studies of such effects must include a combination of extended duration, focused experimentation (feeding, collaring, banding, or full life-cycle testing) to be able to have any success in determining the boundary effects levels, including solid estimates of NOECs and LOECs.

RESPONSE 4-TL-5

Please refer to Response 4-TL-4.

James T. Oris:

It will be necessary and important for the document to address the issue of whether the presentation can be used to support a remediation decision as well as to set clean-up goals. EPA guidance suggests that a good risk assessment be used to do this. I understand that this particular ERA is different than others since it was defined by the consent decree, but this should be highlighted and discussed. It would be useful for a description of the process by which clean-up goals and remediation decisions will be made.

RESPONSE 4-JO-1

More discussion regarding the process to be followed at this site, which is defined in the Consent Decree, will be included in the Regulatory Background section of the Introduction to the revised ERA.

In addition, further clarification on the extent of risk into CT should be addressed. It appears to me that further study downstream of Woods Pond and into CT is needed, and this risk assessment barely addressed the downstream issues.

RESPONSE 4-JO-2

EPA and GE have conducted extensive sampling downstream of the PSA and have evaluated the potential for ecological risk. A summary of the available information and conclusions is presented in the response to General Issue 5.

I think it would be appropriate to describe the issues surrounding the concerns of public commenters for the CT reaches. Namely the extent of the flood plain and the involvement of Native American tribes in the process. It was clear from the questions I asked of EPA that CT agencies were given full opportunity to participate, but did not participate at a level that may have been appropriate. This may have been a political decision on their part, but the attempt to involve them in the process should be noted without being judgmental.

RESPONSE 4-JO-3

Members of CT agencies participated in the historical discussions regarding the development of Assessment and Measurement Endpoints for the ERA. In addition, members of the CT agencies were present during the Consent Decree negotiations, and were provided with multiple opportunities to provide input during the Supplemental Investigation Work Plan.

There are items of revision and clarification that need to be made in the summary discussions and conclusions. For example, the word "catastrophic" is used only in the fish section and not any of the others. It could be argued that none of the receptors are experiencing “catastrophic” failure, so why only point this out for fish? Do we need catastrophic failure in fish populations before we are concerned? It is implicit throughout the document that a HQ>1 is an unacceptable risk; however, this needs to be explicit throughout. The application of risk characterization needs to be clearly defined and applied consistently throughout the document, but currently it is not.

RESPONSE 4-JO-4

EPA agrees; please see response to General Issue 7.

Mary Ann Ottinger:

The summary discussions and conclusions in the ERA are supported by the information presented in the report. These conclusions do describe the risks in an objective, reasonable, and appropriate manner and include consideration of uncertainties in the data. As mentioned in the Introduction, there is a tremendous amount of data that have been collected as part of the ERA by both the EPA and GE teams. The shortcomings pointed out relative to data collected in wildlife do not detract from the wealth of information that is in the ERA and in fact provide the opportunity to address some items to further verify the conclusions of the ERA. In this way, we can be increase the certainty of the conclusions in the ERA and increase the confidence of future actions in the Rest of the River ecosystem.

Brad Sample:

The summary discussions and conclusions were generally suited to the original results presented. However, the final summary and conclusions will need to be revised based on the revisions recommended for specific assessment endpoints outlined above.

The terms acceptable/unacceptable in relation to risks and risk management need to be defined in the context of the assessment, recognizing that in reality these are regulatory/risk management issue and not risk assessment issues.

RESPONSE 4-BS-1

EPA will remove such terms from the revised ERA. Please refer to the response to General Issue 7.

Can this ERA be used to make risk management decisions? Yes, but requires consideration of other issues, such bioaccumulation from abiotic media. In its current form, the much of available analyses are based on contaminant concentrations in biota (especially for the wildlife receptors). Remedial decisions are not generally made based on biota concentrations. As a consequence, additional modeling is needed to relate the biota-based risk conclusions to appropriate soil and sediment concentrations.

RESPONSE 4-BS-2

EPA agrees with the Reviewer's comment, and such additional analyses are anticipated to be performed by GE (and will be by EPA) in the development of the Interim Media Protection Goals (IMPGs) in the process outlined in the Consent Decree.

Are conclusions supported by the data – not exactly in the report as it currently is. However, after revisions in response to panel comments are completed it should.

Ralph G. Stahl:

The summary discussions and conclusions in the ERA are supported by the information in the report. Other than those specific areas identified in my previous comments, the conclusions described the risks in an appropriate manner. There were instances, identified above, where the risks were, in my opinion, not described in an objective nor reasonable manner.

Timothy Thompson:

Benthic Invertebrates

See previous comments.

At this point, the data as presented do not support an MATC that could be used for a risk management decision. However, there are sufficient data, if properly re-evaluated along the lines the Panel recommended, that could form the basis for a supportable MATC.

RESPONSE 4-TT-1

In the revised ERA, EPA will implement the recommendations made by the Reviewers regarding derivation of MATCs. Please refer also to the Response to General Issue 6.

Amphibians

See answer to 4 (j) above.

Fish

See previous comment

Insectivorous Birds

Yes, within the context of specific issues related to statistical analysis, low power associated with the robin field study, and uncertainties, cited above.

Piscivorous Birds

No comment here.

Piscivorous Mammals

Statements such as "significant" and "catastrophic" should be eliminated. Likewise, "unacceptable" is a value judgment, and should not be in the ERA. State the levels of risk, or alternatively, define what is "unacceptable" upfront.

RESPONSE 4-TT-2

Please refer to the response to General Issue 7. There are some cases where the term "significant" is appropriate (e.g., statistical significance evaluations).

As drafted, the ERA needs to consider some revisions and resynthesis. I believe that with the reanalysis of concerns cited by the Peer Review, the ERA will be supportable.

Omnivorous and Carnivorous Mammals

Yes.

5. OTHER PERTINENT INFORMATION

5. To the best of the Panel's knowledge, is there other pertinent information available that was not considered in the ERA? If so, identify the studies or data that could have been considered, the relevance of such studies or data, and how they could have been used in the ERA.

Valery Forbes:

The amount of information considered in this ERA is impressive. I have no knowledge of additional studies or data that could have been considered. Indeed, the copious amounts of data and analyses made reviewing the ERA in a realistic time frame rather challenging. Providing a good roadmap (i.e., Fig 1.1-3) is thus essential. Additional guidance on finding relevant pieces of information would be helpful.

RESPONSE 5-VF-1

Please refer to the response to General Issue 1.

Regarding the overall structure of the ERA, I can appreciate the value of placing the technical details in Appendices for the particularly interested reader. However, I suspect that most readers of this document will either be satisfied with the Executive Summary (perhaps slightly expanded) or will want all of the details. Thus in my view the main chapters provide too much information for the casual reader and too little for the expert.

RESPONSE 5-VF-2

Please refer to the response to General Issue 1.

There were several places in the ERA in which documents 'in preparation' were cited. This should be avoided as a matter of good practice.

RESPONSE 5-VF-3

EPA will update references in the revised ERA for the Modeling Framework Design, which was issued in April 2004.

In the limited time that I have been involved in this project it has been my impression that the cooperation and collaboration between the GE team and the EPA team have been exceptionally positive and constructive. This, I suspect, has not only made the process more pleasant, but I believe has led to a better risk assessment than would have otherwise been done. This ultimately benefits the Housatonic and all of the stakeholder s with interest in it. If it is feasible to formally share the lessons learned in a form that could provide guidance to other risk assessments of this nature, I would strongly encourage that this been done.

RESPONSE 5-VF-4

EPA appreciates the input from the Reviewer and will take advantage, when possible, of opportunities to share such lessons in the appropriate forums.

Thomas W. La Point:

The only consideration I would strongly suggest is to add two things to this ERA: 1) A section not unlike the Executive Summary, but one that includes an overview from the ecological perspective. Although there are to be individual endpoints, some consideration should be given to risk (as for the benthic invertebrates) carried "up" to the fish. This consideration should also take into account the "health" of the population. The Framework for Ecological Risk Assessment focuses on populations. Hence, a discussion of top predators must include "lower tier" assessments of their food webs. It follows that, if the prey organisms are at risk, ultimately, so must the predators be.

RESPONSE 5-TL-1

EPA agrees with the Reviewer's comment that indirect effects, such as effects on food supply, can influence population dynamics. A qualitative discussion of how indirect effects can impact population dynamics was presented in the ERA in Section 12.4.2.2.

2) For the benthic invertebrates, there are several good multivariate approaches that may link the multiple variables to benthic invertebrate abundance and distribution. Such citations include:

Green, R. H. 1979. Sampling Design and Statistical Methods for Environmental Biologists. John Wiley & Sons, NY. 257p.

May, R.M. 1975. Patterns of Species Abundance and Diversity. In M.L. Cody and J.M. Diamond (eds.). Ecology and Evolution of Communities. Pgs. 81-120; Belknap Press, Cambridge, MA.

Moriarty, F. 1999. Ecotoxicology, 3rd Edition. Academic Press, NY.

National Academy of Science. 2001. A Risk-Management Strategy for PCB-Contaminated Sediments. (page 405) National Academy Press, Washington, DC.

Posthuma, L., Suter, G.W., II, and Traas, T.P. 2002. Species Sensitivity Distributions in Ecotoxicology. Lewis Publishers, Boca Raton, FL. (Chapt. 9, pgs. 155 – 193).

Quinn, G.P. and Keough, M.J. 2002. Experimental Design and Data Analysis for Biologists. Cambridge University Press, Cambridge, UK. (Chapter 14, pgs. 380 – 400).

Tabachnick, B.G. and Fidell, L.S. 1996. Using Multivariate Statistics, 3rd Ed. Harper Collins College Publishers, NY. (Chapter 6, pgs. 195 – 238)

Van den Brink, P. J., and Ter Braak, C. J. F, Principal response curves: Analysis of time dependent multivariate responses of biological community to stress. *Environ. Tox. and Chem.*, 18, 138, 1999.

Van Wijngaarden, R. P. A., Van den Brink, P. J., Oude Voshaar, J. H. and Leeuwangh, P., Ordination techniques for analyzing response of biological communities to toxic stress in experimental ecosystems, *Ecotoxicology*, 4, 61, 1995.

RESPONSE 5-TL-2

EPA appreciates the references, and will include additional analyses in the revised ERA. Please refer to the response to General Issue 11.A.

James T. Oris:

Consideration EPA guidance on use of societal value for fish receptors should be included.

RESPONSE 5-JO-1

Please refer to the responses to General Issues 2 and 13.E. Although EPA will include discussion of the DELT information, the revised ERA will differentiate between these individual-level responses and the reproductive measures that have relevance to population endpoints.

Mary Ann Ottinger:

Yes, there is pertinent information that should have been collected, considered, and discussed within the ERA. Although the tree swallow is a commonly studied field species, this species may be a poor choice as a representative species for passerines, despite its presence in the PSA due to the relative insensitivity of this species to PCBs.

RESPONSE 5-MO-1

EPA acknowledged that the tree swallow is an insensitive bird species.

As discussed above, many of the end points chosen for study in wildlife were not particularly sensitive and would reveal impact at a point that the population might be in jeopardy.

RESPONSE 5-MO-2

Please refer to the response to General Issue 14.A.

The statistical models serve to fill in these gaps in that they do provide a conceptual framework in which to make objective determinations. Although these models vary with the underlying assumptions, collapsing the conclusions from the available data and statistical models, using the WOE does provide an overall thoughtful and logical summarization of the risk to the various populations. The ultimate use of these models and the resultant conclusions, using the WOE approach will provide the final demonstration of the degree of responsibility in our stewardship of environmental resources, including the wildlife present within this ecosystem.

Brad Sample:

The data used and evaluated in this ERA is very extensive and comprehensive. I think that due to the combined efforts EPA and GE, virtually al of the relevant data and studies have been identified and incorporated. I am aware of only one study that has recently come out that should be evaluated. This is a 2003 EPA ORD report evaluating avian toxicity data for dioxin-like compounds. The full reference is:

- U.S. EPA. 2003. Analyses of laboratory and field studies of reproductive toxicity in birds exposed to dioxin-like compounds for use in ecological risk assessments. U.S. EPA National Center for Environmental Assessment, Office of Research and Development. NCEA-CIN-1337. 51 pp.

RESPONSE 5-BS-1

EPA will review the cited report and update the revised ERA as appropriate.

Ralph G. Stahl:

In my recommendations and comments above I have identified pertinent information that was not discussed nor presented in detail. It is important that these data be discussed or included in the main text of the ERA or in the Appendices.

RESPONSE 5-RS-1

EPA will evaluate the pertinent information provided by the Reviewer and include it as appropriate in the revised ERA.

Timothy Thompson:

Benthic Invertebrates

There are specific gaps between the ERA Work Plan and the ERA that should be discussed in the ERA. These include:

- Dragonfly surveys – the Supplement Field Investigation Work Plan ("Work Plan) provides in Appendix A.7 a study to identify T&E dragonfly species within the Housatonic River. These species were identified in the PSA, but nothing else was done with that information. The ERA should explain why risks to these species was not identified as being important to the overall ERA.

RESPONSE 5-TT-1

Please refer to the response to General Issue 4.A.

- Freshwater mussels were identified as part of the FSP for the Housatonic River. Appendix A.8 stated that mussels "could potentially be directly or indirectly influenced by PCBs through several mechanisms, including: 1) direct mortality due to accumulation of PCBs in tissues, 2) impaired physiological function, 3) loss of food supplies, 4) loss of fish or amphibian intermediate hosts, and 5) loss or degradation of habitat." Appendix A.15 is designed to identify toxicity to freshwater mussels. While we understand from the presentations on December 18, 2003 to the Panel that these studies did not provide useable data, the ERA should document that finding in the interest of overall transparency.

RESPONSE 5-TT-2

Please refer to the response to General Issue 4.B.

- Crayfish studies were also listed under the Work Plan, yet no data were ever presented and discussed. Appendix A.16 specifically called for the co-location of crayfish samples with the benthic invertebrate samples. There are likely good reasons why these data were not collected or included – but they need to be discussed.

RESPONSE 5-TT-3

Please refer to the response to General Issue 4.D.

- Finally, if there are additional sediment data that have been collected from behind the dams in Connecticut and not currently used in the risk estimates for the ERA, those need to be added to the ERA. The ERA may not conclude that there is no risk to the areas below the PSA, and the ERA should state that the study presentation cannot be a basis for a risk-management decision. The ERA can only qualitatively discuss relative levels of risk outside of the intensively-studies PSA.

RESPONSE 5-TT-4

Please refer to the response to General Issue 5.

Amphibians

- Frequency of TRV exceedances could be estimated for amphibians, and would be helpful in presenting an assessment of risks within the PSA.

RESPONSE 5-TT-5

Please refer to the response to General Issue 1.C.

Fish

- The absence of any reference or incorporation of the biological models – and in particular the diet parameterization is a big whole in the ERA. A considerable amount of work has gone into developing food webs, bioaccumulation and transfer rates, age functions, growth dilution, lipid levels as a function of age, sex, and season, but none of that information is available or used in the ERA. Recommend that these two documents be reconciled before the release of the final ERA.

RESPONSE 5-TT-6

The revised Modeling Framework Design document (WESTON 2004) was in preparation at the time the ERA was produced, but has since been released. Documentation of the calibration effort for the bioaccumulation model is currently in progress. EPA agrees that reference to these parallel studies could contribute to the information in the ERA, and will reference and discuss the studies, as appropriate. This applies to several topics within the fish ERA, including PCB concentration versus percent lipid relationships, selection of representative species, and trophic linkages in the conceptual model.

Insectivorous Birds

No. The field studies and models applied within this study are complete and appropriate.

Piscivorous Birds

No further comments here.

Piscivorous Mammals

No.

Omnivorous and Carnivorous Mammals

I am aware of no other pertinent or relevant information that could be applied here.

ADDITIONAL COMMENTS

Ralph G. Stahl:

Detailed Comments on the ERA

- P 3-72: Please consider providing a fuller explanation for the following statements found in the paragraph from page 3-72, Vol 1 of the ERA.

> "A discussion of attributes considered in the WOE is provided in Section 2, and the rationales for weighting of measurement endpoints are provided in Appendix D. A summary of the derived weightings for each attribute is provided in Table 3.4-1. The chemistry endpoints yielded the lowest overall values because of lower site-specificity and some uncertainties in the biological association between the measurement endpoints and the assessment endpoint(s). The toxicity testing endpoints yielded the highest overall values, because of the high degree of biological relevance of the tests. The benthic community structure endpoints had intermediate values. Although these endpoints were site-specific, collected at a time when effects would be expected, and were measures of the community structure component of the assessment endpoint, the potential for the confounding effects of other factors in the direct attribution of the response to the stressor reduced the utility of these endpoints to some degree."

RESPONSE A-RS-1

Please refer to the response to General Issue 8.A.

- P3-75, Table 3-4.2: Despite the widespread use of the WOE process, it would be useful to more fully explain how the finding of a chemical in a medium can be evidence of harm or impact when there have been no measures of biological response. This may be more an issue of semantics than of substance.

"C1 – concentration in the water column"

"C2- concentration in the sediment"

RESPONSE A-RS-2

EPA will provide clearer definitions in the revised ERA.

- P 3-78. How does an LC50 or EC50 indicate an ecologically significant response? These values are generally relevant to acute toxic responses but those are not equivalent to ecologically significant responses. Unless these apply to ecologically relevant endpoints such as reproduction, then it is inaccurate to suggest they have an ecological relevance.

> "Above a concentration of 3 mg/kg tPCB, numerous endpoints indicated ecologically significant responses, with many LC50/EC50 values falling in this range."

RESPONSE A-RS-3

EPA agrees that 50% effects levels, whether for lethal or sublethal endpoints, are not equivalent to threshold levels for ecologically significant responses. It is for this reason that threshold concentrations for different effect sizes (i.e., EC20 and EC50) were calculated. However, most of the endpoints referenced on page 3-78 are either survival or reproduction endpoints, and 50% reductions of the metrics for these endpoints are widely recognized as exceeding the level at which population-level responses are observed.

- P 3-79. As noted in an earlier comment, the designation of a risk being "unacceptable" is not a question answered by an ERA. It is a risk management decision as to whether the finding of a risk is unacceptable. At best the ERA is charged with evaluating the risks, quantifying them if possible, and indicating the sources of those risks.

"Unacceptable risks are predicted for the majority of sediment sampled within Reach 5A."

RESPONSE A-RS-4

Please refer to the response to General Issue 7.

- Section 4: Amphibians. As a general comment on this section, I suggest the authors of the ERA consult the recently published book "*Amphibian Decline: An integrated analysis of multiple stressor effects*". Linder, G., Krest, S.K. and Sparling, D.W. (eds) 2003, SETAC Press, Pensacola, FL. I believe it will help in the interpretation of both the leopard and wood frog data sets.

RESPONSE A-RS-5

EPA agrees that this publication provides useful information for interpreting the amphibian data and, in fact, used this information in the interpretations made in the ERA. Dr. Fort, EPA's Principal Investigator for the wood frog and leopard frog studies, contributed a chapter to this book. The authors of the amphibian section of the ERA have read the book cited by the Reviewer, and have also recently taught a short-course with G. Linder and S. Krest at SETAC on the use of amphibians in risk assessment.

- P 4-34. Please explain how conclusions can be drawn that implicate the presence of PCBs in soils/sediments as related to presence or absence of leopard frogs in the contaminated areas when no leopard frogs were found in the reference locations.

"4.4.1.2.1 Leopard Frog Study Reproductive Fitness: Adult male and female leopard frogs (and some juveniles of both sexes) were collected from the nine contaminated sampling areas in the PSA and transported to Fort Environmental Laboratories, Inc. (FEL). No leopard frogs were collected at the three reference areas; therefore, control animals purchased from a commercial supplier (Carolina Biological Supply, CBS) were used. These frogs were collected in Vermont directly upon order, shipped to CBS, and then forwarded to FEL (formerly part of The Stover Group)."

RESPONSE A-RS-6

Leopard frogs were observed in reference areas in previous years, but were not present in sufficient numbers at the time of collection for the leopard frog study. The conclusions regarding the effects of PCBs to leopard frogs were based on other lines of evidence. Please refer to the response to General Issue 12.A.

- P 4-36, Figures 4.4-1, 4.4-2. These figures would, without benefit of the initial text on the leopard frog study, lead the reader to believe that adult frogs were obtained from the "reference areas". These figures are misleading in that respect and a notation should be added to them to remind the reader that no adult frogs were collected from the reference areas.

RESPONSE A-RS-7

EPA agrees and will modify the labels on the Figures 4.4-1 and 4.4-2 in the revised ERA.

- P 4-36, Figure 4.4-2. The percentage of mature Stage IV oocytes in adult female frogs from the so-called "reference area" seems to indicate a particularly striking impact on oocyte maturation at very low PCB concentrations as indicated by the vertical axis. Even in areas with extremely low but measurable levels of PCBs in the sediments, there are few if any mature stage IV oocytes. What is the basis for the last sentence in the paragraph below when there is no correlation between sediment PCB concentration and depression of mature Stage IV oocytes?

"Few of the PSA sites produced female specimens that possessed any biologically significant quantity of Stage VI oocytes (mature eggs capable of fertilization), with the exception of Station W-7a (Figure 4.4-2). Immature oocytes (< Stage III) were observed in mature female specimens collected from all PSA sampling areas, however developing oocytes were found in specimens from Sites W-7a, W-4, EW-3, and W-1 (18.0, 0.5, 30.0, and 0.2 mg/kg sediment tPCBs, respectively). Therefore, the lack of success in artificially fertilizing oocytes from contaminated site specimens was not surprising, and appeared to be the primary limiting factor in the reproductive dysfunction observed in the contaminated site specimens evaluated from the PSA."

RESPONSE A-RS-8

The lack of success in artificially fertilizing oocytes was believed to be due to the delayed onset of their development. As explained in Response 3.2-VF-3, female leopard frogs were collected at a time when they should have had mature oocytes. This expectation is supported by the fact that other leopard frogs were laying eggs at the same time in 2000, and also in 2003, and that other metrics indicating breeding readiness, such as vocalizing males, were observed consistently throughout the period of collection. In the revised ERA, additional text will be added explaining why EPA and Dr. Fort agree that there appears to be a threshold level effect to leopard frogs at very low PCB concentrations, which resulted in delayed oocyte development.

- P 4-37. Similar to comments noted above for P 4-36, please clarify, what adult male frogs are being referenced in Table 4.4-1 with regards to sperm head abnormalities. Based on the description in the text of this table, the "reference area" frogs were in fact purchased from Carolina Biological Supply and did not originate from the Housatonic River watershed.

RESPONSE A-RS-9

The reference frogs referred to in Table 4.4-1 were the control leopard frogs. The revised ERA will contain text clarifying why these frogs are suitable for this comparison.

- P 4-44. Please clarify the statements below concerning conclusions of the leopard frog study. At the station (E-1) with the highest level of tPCBs in the sediment (E-1, 160 mg/kg), the mean percent abnormal sperm head was 14.3. Even with the data manipulations (spatial weighting for sediments), and the presence of tPCBs in the water column, there would appear be a very high uncertainty associated with the conclusions detailed in the text box. There appears to be an absence of a concentration response for both the male and female data, further complicated by the inability to judge the findings against reference areas since no adult male or female frogs were found in them.

> "male and female adult frogs showed signs of reproductive stress, with the females showing more severe effects. Males exhibited a high incidence of malformed sperm in the higher-sediment tPCB sites (up to 50%). Females had virtually no mature eggs (Stage VI, which the eggs must reach in order for fertilization to occur). Incidences of immature oocytes (Stage III or smaller) were high in the sites with high concentrations of sediment tPCB (up to 99% Stage III)."

RESPONSE A-RS-10

The lack of leopard frogs in the reference areas was addressed in the ERA in Appendix E, Section 5.1 The statements quoted from the ERA as part of this comment were developed based on a careful evaluation of all of the data presented in the risk assessment. This includes data generated by EPA, by GE, and by other researchers as published in the literature.

Some of the sample sizes in the leopard frog study precluded the development of a dose-response relationship. This was recognized in the ERA. Wood frog sample sizes, however, were sufficiently robust to allow the development of dose-response relationships that were statistically significant for numerous measurement endpoints, including malformations and sex skewness. There were also observed effects that FEL attributed to both PCB concentrations in tissue and sediment and concentrations published in the literature, as reported in Appendix E, Section 3.6.1.

A review of the extent of PCB contamination in the floodplain, in comparison to concentrations observed to have effects in FEL's studies and in the literature, suggests that amphibian populations are impacted throughout much of the PSA. Text in the revised ERA will be edited to clarify these observations, in keeping with other Reviewer comments.

- P 4-46. Table 4-45. It is difficult to determine the basis for the "n" in this table. Please clarify what it represents. For example, does each "n" represent an individual frog assessed for the various endpoints over time, or different frogs collected over time for the specific endpoint measured, or something else all together. This is not clear.

RESPONSE A-RS-11

The "n" represents the number of samples being compared. For example, in the first line of the table, the "11" refers to 11 tPCB sediment samples and 11 tPCB egg mass samples. In the revised ERA, a note will be added to the table to clarify this column.

- P 4-67. Table 4-5-1. It is difficult to recall the literature based toxicity thresholds (1 and 10 mg/kg) while reviewing this table. It would be helpful for the reader if the toxicity threshold were noted either in the title of the table, on the table, or in a footnote to the table. More specifically, it would be important to clearly identify which of the thresholds was used in the comparison.

RESPONSE A-RS-12

EPA agrees that the table can be improved by showing the toxicity threshold values on Table 4.5-1 in the revised ERA.

- P 4-78. Please add units to the PCB concentrations cited in the paragraph below.

"The Phase III wood frog metamorph results for the Housatonic sampling exhibited a range of sex ratios. Sediment in the two most contaminated vernal pools (38-VP-1 and 38-VP-2) contained PCB concentrations of 28.5 and 32.3 PCBs based on spatial weighting of sediment data."

RESPONSE A-RS-13

The units are mg/kg. The units will be added to the text in the revised ERA.

- P 4-78. The term "feminization" is used incorrectly in the paragraph below. Typically it refers to the alteration in males whereby morphologically or in their behavior they are essentially female. It does not mean a change in the sex ratio of a feral population of organisms. If there are morphological or behavioral changes in the frogs studied, it is not made clear in the text.

"The Housatonic River vernal pool sex ratio data for wood frog metamorphs and breeding adults exhibit strong differences from Berven's data at a non-contaminated site. The general trend for the wood frogs examined near the Housatonic River PSA is a marked decrease in the male to female ratio in both metamorphs and breeding adults. This feminization of the wood frogs in this study may be adversely impacting the local population."

RESPONSE A-RS-14

The term "feminization," as used in the text, does not refer to changes in male behavior or morphology, it refers to an increase in the total number of females observed (i.e., sex ratio). Text will be added to the revised ERA to clarify this point.

- P 4-82. Given that no adult leopard frogs were found in the "reference areas", and the range of PCB concentrations in sediments was not strongly correlated with some of the endpoints measured, how is the statement below fully supported by these highly variable results ? There may be validity to the statement with respect to the wood frog, but this would not appear to be fully supported by the leopard frog results.

"ERA findings suggest that amphibian populations are impacted throughout much of the PSA, with leopard frogs impacted at a wide range of sediment concentrations (likely due to the life history of contact with sediment PCB concentrations, which were not measured in the study), and with stronger responses from wood frogs expected in the more highly contaminated vernal pools. The indications of community responses from the population studies (i.e., localized depressions of richness and abundance near high tPCB vernal pools, and high incidence of malformations observed) substantiate these findings."

RESPONSE A-RS-15

The lack of leopard frogs in the reference areas was addressed in the ERA in Appendix E, Section 5.1. The statements quoted from the ERA as part of this comment were developed based on a careful evaluation of all of the data presented in the risk assessment. This includes data generated by EPA, by GE, and by other researchers as published in the literature.

Some of the sample sizes in the leopard frog study precluded the development of a dose-response relationship. This was recognized in the ERA. Wood frog sample sizes, however, were sufficiently robust to allow the development of dose-response relationships that were statistically significant for several measurement endpoints, including malformations and sex skewness. There were also observed effects that FEL attributed to both PCB concentrations in tissue and sediment and concentrations published in the literature, as reported in Appendix E, Section 3.6.1.

A review of the extent of PCB contamination in the floodplain, in comparison to concentrations observed to have effects in FEL's studies and in the literature, suggests that amphibian populations are impacted throughout much of the PSA. Text in the revised ERA will be edited to clarify these observations, in keeping with other Reviewer comments.

- P 5-17. Please explain the scientific reasoning for the statements below. Although I have not read the cited paper, it is difficult to understand how benzo(g,h,i)perylene could be as metabolically active in fish as pyrene for example. Thus the statement that PAHs appear to be rapidly metabolized and therefore fish tissues were not analyzed for these constituents would seem to be a substantial oversimplification of the issue.

"5.2.3 Sediment Chemistry Assessment (Exposure to PAHs)

"There were no data on fish tissue concentrations in the PSA for the eight individual PAHs retained as COCs for fish, or for total PAHs because PAHs are readily metabolized by most aquatic biota, including fish (Johnson 2000). Exposure for these contaminants was therefore assessed based on sediment concentrations only."

RESPONSE A-RS-16

Many polycyclic aromatic hydrocarbons (PAHs) are capable of inducing the Mixed Function Oxygenase (MFO) system. The MFO system is responsible for the rapid metabolism and excretion of PAHs, and therefore limits bioaccumulation. However, exposure to PAHs has been linked to reproductive impairment, immune dysfunction, increased incidence of liver lesions, and other histopathological endpoints in fish (Malins et al. 1987; Johnson et al. 1988; Varanasi et al. 1992). Metabolism of some PAHs creates negative health impacts via toxic intermediates and daughter products.

Because fish rapidly metabolize and excrete PAHs, fish tissue residue concentrations of parent PAH compounds do not provide a useful measure of exposure for fish (Field 1996; Varanasi et al. 1989). EPA did not mean to imply that PAHs did not merit consideration due to the existence of a metabolic pathway. Instead, EPA stated that measurement of parent PAH compounds in fish tissues does not provide a reliable means of evaluating risk. Although metabolism varies among PAH compounds, it is sufficiently strong to preclude effective exposure analysis based on the concentrations of parent compounds alone.

- P 5-19. Please explain what percentage of the samples collected for Table 5-24 were actually a calculated concentration stemming from the use of the detection limit (DL) rather than an actual analytical tissue concentration.

RESPONSE A-RS-17

Table 5-24 was included in Section 5 as a summary of the detailed material presented in Appendix F. Tables F.2-9 through F.2-13 in Appendix F present the results of a bounding analysis performed with TEQ concentrations, using both ND=DL and ND=0. EPA believes that the bounds of the TEQ concentrations, for a variety of statistical measures (minimum, maximum, mean, median, percentiles), provide a more comprehensive and useful evaluation of the data, as opposed to presenting only the detection frequencies. The labeling of Table 5-24 will be reviewed and clarified if necessary.

- P 5-20. For Table 5-25 please explain the basis for the differences in the number of sediment samples collected for PAH analysis among the various reaches. It is not clear from the text or the table why there is such a large difference.

RESPONSE A-RS-18

The difference in sample sizes is related to the length of river channel associated with each reach (Figure 1.4-1 of ERA). Reaches 5A and 5C each consist of approximately 4 miles of river channel, and have the largest number of sediment PAH samples (20 and 11, respectively). Reach 5B and Woods Pond (including headwaters) are only about 2.5 miles and 0.4 miles in length, respectively, and therefore have only 6 and 3 samples, respectively.

- P 5-32. What “extracts” are referred to in the statements below ? Water, sediment, other ? The reference to “extracts” occurs several times in this section and in the following pages, yet there is no specification as to what extract was tested. Please clarify.

“survival was also observed in medaka exposed to Housatonic River extracts. Between 3 and 15 days post swim-up, medaka exposed to extracts from Reach 5BC and Reach 6 exhibited statistically significant reductions in survival relative to control fry. Survival was not affected in largemouth bass and rainbow trout exposed to Housatonic River extracts.”

RESPONSE A-RS-19

The term “extracts” refers to tissue extracts (i.e., organic contaminants extracted from largemouth bass tissue collected in the Phase I studies) administered to control largemouth bass, medaka, and rainbow trout eggs exposed via injection. In Phase II of the fish toxicity studies conducted by USGS, organochlorine contaminants present in largemouth bass from the Phase I study were extracted and injected into cultured eggs of these three species.

For more detail, please refer to Section F.3.3.2.1 of the ERA.

- P 5-33. Please discuss how the influence of handling the eggs and the potential synergistic effects of damaging the physical integrity of the egg (via injection) is not also one of the potential confounding influences on the outcome of these studies ? It is generally the case, in my experience, that rainbow trout and medaka eggs, while appearing to be hardy during laboratory manipulation, are in fact quite sensitive to physical trauma as well as that associated with exposure to toxic substances (and to direct light in some cases). Often this sensitivity is not evidenced until several days or weeks after fertilization and handling. Dr. Tillet, in his personal communication with the authors of the ERA, is no doubt aware of this and may have shed light on it, but this is not noted in this particular section. This is, in my opinion, shown clearly in Figure 5.3-3 where the incidence of deformities in uninjected eggs and those injected with triolein. Please explain.

“Overall, medaka at 15 days post swim-up exhibited the lowest LD50s, relative to other species, for all extracts and standards, with the exception of TCDD. The overall results (i.e., order of magnitude difference in TEQ-based LD50s between site extracts and standards) appears to indicate that the Housatonic River extracts are more toxic than would be predicted on the basis of an additive model of dioxin-like toxicity alone. The increased toxicity observed with the Housatonic River extracts could be attributed to synergistic effects of PCB mixtures and effects of other PCBs in the mixture that are not considered using the TEQ approach (Tillitt, personal communication 2003).”

RESPONSE A-RS-20

EPA will provide further discussion of the potential influence of handling stress (and/or other factors impacting control performance) in the revised ERA. However, Dr. Tillitt, the Principal Investigator, does not believe that handling stress was a concern in the study. Please refer to the response to General Issue 13.A.

- P 5-34. One could argue that these were not exposures "in ovo" per se, but most specifically exposure via injection. In ovo would connotate exposure of the entire egg to the substance of concern in an aqueous exposure scenario, without penetration of the egg membrane. Based on the description in the text, the latter was not done. I understand the purpose of these experiments but I take issue with the use of the terminology "in ovo".

"Fish exposed in ovo to high doses of Housatonic River extracts exhibited similar types of gross pathologies as the dioxin-like standards, including craniofacial deformities, fin deformities, spinal deformities, swim bladder deformities, hemorrhage, pericardial edema, peritoneal edema, …"

RESPONSE A-RS-21

EPA and the Principal Investigators conducting this study believe that the term "in ovo" is applicable to toxicity experiments in which a chemical is injected into an embryo. This is supported by numerous published studies. For example, Rodríguez-Burgos (2003) states that "some in ovo experiments involved injection of its heterologous antiserum from rabbit into the yolk sac." The term in ovo was also used for recent published contaminant injection studies using fish (Papoulias et al. 2003).

Although EPA believes the term is appropriately applied, EPA will use more specific terminology in the revised ERA.

- P 5-34. One could argue that another interpretation of the lack of a dose response, and that similar deformities were observed in fish when eggs were injected with the "extracts" from the PSA and reference areas, as well as the control, is that the experimental design was inappropriate. Please explain.

"Some of the deformities observed in fish were only weakly related to tPCB or TEQ concentrations for one species/life stage/treatment combination. The lack of a dose-response in fish injected with Housatonic River extracts and/or PCB and TCDD standards and the occurrence of these deformities in fish injected with control and reference site extracts indicates that these abnormalities are not the most reliable markers of PCB exposure."

RESPONSE A-RS-22

Please refer to the response for General Issue 13.A. EPA acknowledges that not all trials or deformity types exhibited a linear or sigmoidal concentration-response relationship, nor would such responses be expected consistently in a study of this type. Certain responses to PCBs are not expected to be well predicted with a simple model. The responsiveness of any endpoint of chemical effect is related to the level of biological organization (molecular, biochemical, histological, etc.) and the proximity of the response to the known mode of action of the chemical or class of chemical. This is due to the greater number of mitigating factors involved in the response at each increase in biological level of organization. Therefore, it is not expected that a simple model would predict the higher-level responses with the same level of precision or accuracy as lower level responses.

Given the lack of precise understanding of the mechanism of PCB toxicosis at the whole organism level, individual deformities measured are not all expected to exhibit a dose-dependent response. This is not indicative of a flaw in experimental design, but rather is an expression of the inherent uncertainty of the experiment. Rather than select a subset of the pathologies for statistical analysis, which would have potential for bias, both the principal investigator and EPA evaluated the cumulative incidence of all major pathologies. The aggregation of response types, combined with the use of control corrections, protects against spurious statistical outcomes. The presence of baseline rates of deformities in control fish, and/or the observation of select individual pathologies that do not exhibit dose-response, do not provide a basis for concluding either that PCBs do not cause biological effects or that the experiments were improperly designed to measure effects.

- P 5-38. The discussion of the cytochrome P450 studies appears to be overly simplified and not particularly informative. I am at a loss to understand what exactly is meant by "tissues"(liver ?) since the text tends to mix references to the egg injection work, and P450 induction in tissue. Please clarify.

"Cytochrome P450

"Cytochrome P450 induction was evaluated qualitatively in largemouth bass, medaka, and rainbow trout tissues using immunochemical histological techniques. Cytochrome P450 induction was observed in fish exposed to both standards and Housatonic River extracts. Rainbow trout was the most sensitive test species, exhibiting apparent dose-related increases in cytochrome P450 induction. The strongest response (i.e., highest induction) was observed in trout exposed to Reach 5BC extracts. Low and moderate level cytochrome P450 induction was observed in bass exposed to 6 µg TCDD/kg egg and medaka exposed to 2 to 6 µg TCDD/kg egg. Medaka also exhibited moderate dose-related cytochrome P450 induction following exposure to reference site extracts containing 0.15 mg tPCBs/kg egg. Largemouth bass did not appear to exhibit a dose-related induction of cytochrome P450 following exposure to Housatonic River extracts."

RESPONSE A-RS-23

Please refer to the response to General Issue 13.A. A more complete presentation of endocrine, histological, and enzymatic endpoint responses will be provided in the revised ERA based upon the Reviewers' comments. The information is currently presented in the report prepared by Tillitt et al. (2003).

- P 5-38. Ibid, earlier comments concerning confounding influences on the egg injection study.

"The increased toxicity associated with the Housatonic River extracts could be attributed to synergistic toxicity of the PCB mixtures, as well as the effects of PCBs that are not incorporated into the TEQ model (Tillitt, personal communication 2003)."

RESPONSE A-RS-24

Please refer to the response to General Issue 13.A, and Responses A-RS-20 and A-RS-22, above.

- P 5-39, 5-40. I believe the ED concept is not used or defined appropriately by the authors as there has been a mixing of sublethal and lethal effects. In classical toxicology the ED is that dose which elicits a non-lethal response whereas the LD is that dose which elicits a lethal response. It is inappropriate, unless it is made clear that this definition is specific to this study alone, to use the term in this manner. Further, as I stated in previous comments and questions, I do not believe it is the purpose of the ERA to make judgments as to the acceptability or unacceptability of a particular level of harm or risk.

"5.3.3.2.1 ED50 Estimates

"ED50s derived from the Phase II study data were used to develop thresholds for Housatonic River extracts and PCB-126 and TCDD standards. An ED50 value represents the concentration at which sublethal or lethal effects (i.e., either mortality or one or more abnormalities) was observed in 50% of the population, relative to the negative controls. This combined toxicity endpoint provides an indication of the concentration threshold for sublethal and lethal effects in early life stages of fish. The ED50 endpoint represents a large effect size and indicates an unacceptable level of biological harm."

RESPONSE A-RS-25

EPA will provide additional clarification in the revised ERA on the use of the term ED to represent combined survival and major pathologies. As indicated in the response to General Issue 7, EPA agrees with the Reviewer's comment regarding value-laden terminology and will revise the ERA accordingly.

- P 5-42. Figure 5.3-8. All labels for the x-axis in this figure are illegible. All labels have been blacked out in the current figure.

RESPONSE A-RS-26

This labeling problem is apparently the result of an incompatibility between software systems used to generate the information for the report. The labels will be corrected in the revised ERA.

- P 5-44. It is not clear that the exclusion of the reference station is justified based on the short discussion in the paragraph below. Does the exclusion of data illustrating the lack of response at lower concentrations force the MATC downwards ? Why not incorporate all the relevant information into the MATC estimation ?

"Exclusion of Reference Station—Threemile pond extracts were excluded from the MATC derivation; the maximum PCB concentrations tested for these reference fish (0.15 mg/kg tPCB and 6 pg/g TEQ) were insufficient to yield large toxic responses or provide meaningful information on the magnitude of the ED50 value. The concentrations in these extracts were well below the levels causing effects in the contaminated site extracts."

RESPONSE A-RS-27

The exclusion of data indicating a lack of response at lower concentrations does not force the MATC downward. On the contrary, the data were excluded to avoid derivation of an MATC that is artificially lowered due to inclusion of unbounded no-effect concentrations. Because exposure concentrations in the reference treatment were far below levels exhibiting adverse responses, there is no way to meaningfully incorporate the unbounded no-effect concentrations in the MATC derivation.

- P 5-55. What is the definition of "ecologically significant" in the context of the discussion shown below ? It is not clear that all of the endpoints or results used in the evaluation were those which might be considered as ecologically relevant.

"Figure 5.4-1 depicts hazard quotients for PSA fish tissue concentrations compared to the average of the site-specific (Phase I and Phase II) fish effects thresholds derived for the PSA (i.e., 49 4 mg/kg tPCB). All mean HQs are below 3 and median HQs are below 2, indicative of an ecologically significant but low magnitude of risk."

RESPONSE A-RS-28

Please refer to the responses to General Issue 7, and Response 4-TT-2.

- P 5-63. Table 5.4-2. Please explain the designation "DL" in the table. It would be helpful to note its meaning on the table itself.

RESPONSE A-RS-29

The term "DL substitution" applies to calculations for which non-detected values were replaced with the detection limit. EPA will note the definition of this term (and also the term "0 substitution") in the revised ERA.

- P 5-64. Please explain the scientific basis, other than the application of professional judgment, for the statement that an HQ of greater than 1 but less than 3 is indicative of "ecologically significant", but low magnitude risk.

"Figure 5.4-7 depicts hazard quotients for PSA fish tissue concentrations compared to the average site-specific effects threshold (i.e., 42 ng/kg TEQ). All 75th percentile-based HQs exceed 1, but mean and median HQs for adult fish are below 3 for all species. These HQs are indicative of ecologically significant but low magnitude risk."

RESPONSE A-RS-30

Please refer to the response to General Issue 7. The assignment of risk magnitude for fish was made in the weight-of-evidence (WOE) process, not by converting an HQ magnitude to a risk conclusion. The median HQ values slightly above 1 are consistent with, but were not used as the basis for, the determination of "low risk." EPA agrees that the text on page 5-64 is ambiguous and will correct it in the revised ERA.

- P 5-74. The sentence below is confusing and requires editing. I'm not sure in what context rainbow trout would be more "toxic".

"Although the ED50 values for trout were within a factor of 2 of warmwater species in the Phase II trials, other indications of toxicity (Tillitt et al. 2003b) suggest that rainbow trout were slightly more toxic than the warmwater species."

RESPONSE A-RS-31

EPA concurs with the comment. The term "more toxic" will be rephrased to "more toxicologically sensitive" in the revised ERA.

- P 5-74. Please explain the scientific basis for the use of a factor of "4" to adjust the MATC.

"Furthermore, the rainbow trout strain applied in the Phase II testing (Tillitt, personal communication 2003) is less sensitive than other test strains, and the sensitivity of other downstream trout species (e.g., brown trout) has not been assessed. Therefore, the PSA effects threshold of 49 mg/kg tPCB was divided by a factor of 4 to account for potential increased sensitivity of downstream coldwater species (i.e., coldwater MATC of 12 mg/kg tPCB whole body, wet weight)."

RESPONSE A-RS-32

Please refer to the response to General Issue 13.B.

- P 9-44. Were these studies conducted under the applicable Good Laboratory Practice (GLPs) regulations ?

"All procedures were executed according to CERC Standard Operating Procedures (SOPs) and QA/QC procedures."

RESPONSE A-RS-33

The mink enzyme study followed procedures consistent with GLP. Details on the procedures are provided in the Principal Investigators' report. QA/QC procedures were as stringent, if not more stringent, than established GLP procedures.

- P 10-32. Were there no data from studies in dogs such that those data could be used as a surrogate for the red fox ? Using the rat as a surrogate for red fox would not appear to be scientifically supportable given the substantial differences in metabolism, etc.

> "As a result, laboratory studies involving surrogate species were used to estimate effects to the representative species. For short -tailed shrew and red fox, the rat was used as a surrogate species for effects due to exposure to tPCBs. In the case of exposure to TEQ, the mouse was used as a surrogate species for short-tailed shrew, while the rat was used for red fox."

RESPONSE A-RS-34

Please refer to Response 3.3-RS-6.

References

Barnthouse L. W., Glaser D., & Young J. (2003) Effects of historic PCB exposures on the reproductive success of the Hudson River striped bass population. *Environmental Science and Technology* 37: 223228.

Diamond J. M. & Serveiss V. B. (2001) Identifying sources of stress to native aquatic fauna using a watershed ecological risk assessment framework. *Environmental Science and Technology* 35: 47114718.

Drevnick P. E. & Sandheinrich M. B. (2003) Effects of dietary methylmercury on reproductive endocrinology of fathead minnows. *Environmental Science and Technology* 37: 4390-4396.

Fox G. A. (1991) Practical causal inference for ecoepidemiologists. *Journal of Toxicology and Environmental Health.* 33: 359-373.

Pittinger C. A., Turnley J. G., & Mehrle P. M. (2001) A multistakeholder ecological risk management framework. In: *Risk Management: Ecological risk-based decision making.* (eds R. G. Stahl, R. Bachman, A. L. Barton, J. R. Clark, P. L. deFur, S. J. Ells, C. A. Pittinger, M. W. Slimak, and R. S. Wentsel) pp. 2132. SETAC Press, Pensacola, Florida.

Stahl R. G., Bachman R., Barton A. L., Clark J. R., deFur P. L., Ells S. J., Pittinger C. A., Slimak M. W., & Wentsel R. S. (2001) Risk Management: Ecological Risk-Based Decision Making. SETAC Press, Pensacola.

USEPA (2000) Stressor Identification Guidance Document. USEPA, Washington, DC.

REFERENCES

ASTM (American Society for Testing and Materials). 2001. Biological Effects and Environmental Fate; Biotechnology; Pesticides. A2. Guidance for conducting sediment toxicity tests with *Daphnia magna* (*D. magna*) and *Ceriodaphnia dubia* (*C. dubia*). E1706. In: *Annual Book of ASTM Standards*, Vol. 11.05, Philadelphia, PA.

ASTM (American Society for Testing and Materials). 2004. *Standard Guide for Conducting Early Life-Stage Toxicity Tests with Fishes*. E1241-98.

Bailer, A.J. and J.T. Oris. 1997. Estimating inhibition concentrations for different response scales using generalized linear models. *Environmental Toxicology and Chemistry*. 16:1554-1559 pp.

Barnthouse, L. W., D. Glaser, and J. Young. 2003. Effects of historic PCB exposures on the reproductive success of the Hudson River striped bass population. *Environmental Science and Technology* 37: 223228.

BBL (Blasland, Bouck & Lee, Inc.). 2002. *Evaluation of Mink - Presence/Absence, Distribution, and Abundance in the Housatonic River Floodplain*. Report prepared for the General Electric Company.

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

Beatriz, J.M., M.L. Rosenshield, and W.H. Karasov. 2000. Effects of PCB 126 and ammonia, alone and in combination, on green frog (*Rana clamitans*) and leopard frog (*R. pipiens*) hatching success, development, and metamorphosis. *Journal of the Iowa Academy of Science*. 107:113-122.

Beckett, K.J., S. O'Reilly, B. Yamini, J. McCormick and S.J. Bursian. 2003. Can the mandibular and maxillary squamous cell proliferation in mink (*Mustela vison*) be considered neoplastic? 24th Annual Society of Environmental Toxicology and Chemistry Meeting, Austin, Texas.

Bleavins, M.R., R.J. Aulerich, and R.K. Ringer. 1980. Polychlorinated Biphenyls (Aroclors 1016 and 1242): Effects on survival and reproduction in mink and ferrets. *Archives of Environmental Contamination and Toxicology* 9:627-635.

Boonstra, Rudy. 2003. Attachment R, Comments Related to Short-Tailed Shrew Demography Study. In: *Comments of General Electric Company on the Ecological Risk Assessment for the General Electric/Housatonic River Site, Rest of River (July 2003 Draft).* Prepared for General Electric Company.

Branton Environmental Consulting and BBL Sciences. 2003. Attachment N, Analysis of Kit Survivability Data from EPA's Mink Feeding Study. In: *Comments of General Electric Company on the Ecological Risk Assessment for the General Electric/Housatonic River Site, Rest of River (July 2003 Draft).* Prepared for General Electric Company.

Bursian, S.J., R.J. Aulerich, B. Yamini, and D.E. Tillitt. 2002. *Dietary Exposure of Mink to Fish from the Housatonic River: Effects on Reproduction and Survival.* Michigan State University, Department of Animal Science. Final Draft Report. March 28, 2002.

Bursian, S.J. and B. Yamini. 2003. *Dietary Exposure of Mink to Fish from the Housatonic River. Inducement of Manidibular and Maxillary Squamos Cell Proliferation*. Report submitted to Weston Solutions, Inc. March 18, 2003.

Burton, G.A., B.L. Stemmer, K.L. Winks, P.E. Ross, and L.C. Burnett. 1989. A multitrophic level evaluation of sediment toxicity in Waukegan and Indiana Harbors. *Environmental Toxicology and Chemistry*. 8:1057-1066.

Burton, Jr, G.A., G.E. Batley, P.M. Chapman, V.E. Forbes, E.P. Smith, T. Reynoldson, C.E. Schlekat, P.J. denBesten, A.J. Bailer, A.S. Green, and R.L. Dwyer. 2002. A weight-of-evidence framework for assessing sediment (or other) contamination: improving certainty in the decision-making process. *Human and Ecological Risk Assessment*. 8(7):1641-1655.

Chadwick (Chadwick & Associates). 1994. *Aquatic Ecology Assessment of the Housatonic River, Massachusetts*. Prepared for General Electric Company, Pittsfield, MA.

Chapman, P.M. and G.S. Mann. 1999. Sediment quality values (SQVs) and ecological risk assessment (ERA). *Marine Pollution Bulletin*. 38(5):339-344

Chapman, P.M., B.G. McDonald, and G.S. Lawrence. 2002. Weight of evidence issues and frameworks for sediment quality (and other) assessments. *Human and Ecological Risk Assessment*. 8(7):1489-1515.

Custer, C.M. 2002. *Exposure and Effects of Chemical Contaminants on Tree Swallows Nesting Along the Housatonic River, Berkshire Co., Massachusetts, 1998-2000*. Final report to U.S. Environmental Protection Agency. USGS, Upper Midwest Environmental Sciences Center, La Crosse, WI.

DeWitt, T.H., R.J. Ozretich, R.C. Swartz, J.O. Lamberson, D.W. Schults, G.R. Ditsworth, J.K.P. Jones, L. Hoselton, and L.M. Smith. 1992. The influence of organic matter quality on the toxicity and partitioning of sediment-associated fluoranthene. *Environmental Toxicology and Chemistry*. 11:197-208.

Diamond, J. M. and V. B. Serveiss. 2001. Identifying sources of stress to native aquatic fauna using a watershed ecological risk assessment framework. *Environmental Science and Technology* 35: 4711 - 4718.

DiToro, D.M., C.S. Zarba, D.J. Hansen, W.J. Berry, R.C. Swartz, C.E. Cowan, S.P. Pavlou, H.E. Allen, N.A. Thomas, and P.R. Paquin. 1991. Technical basis for establishing sediment quality criteria for nonionic organic chemicals using equilibrium partitioning. *Environmental Toxicology and Chemistry*. 10:1541-1583.

DiToro, D.M., and D. DeRosa. 1998. Equilibrium Partitioning and Organic Carbon Normalization. Proceedings, National Sediment Bioaccumulation Conference, September 11-13, 1996, Bethesda Maryland. Published by United States Environmental Protection Agency, Office of Water, EPA 823-R-98-002. February 1998. pp. 3-3 to 3-6.

Drevnick P. E., and M.B. Sandheinrich. 2003. Effects of dietary methylmercury on reproductive endocrinology of fathead minnows. *Environmental Science and Technology* 37: 4390-4396.

Efroymson, R.A., G.W. Suter, B.E. Sample, and D.S. Jones. 1997. *Preliminary Remedial Goals for Ecological Endpoints*. Oak Ridge National Laboratory for U.S. DOE. ES/ER/TM-162/R2. August 1997.

Eisler, R. 1987. Polycyclic Aromatic Hydrocarbon Hazards to Fish, Wildlife, and Invertebrates: A Synoptic Review. *Contaminant Hazard Reviews, Biological Report* 85(1.11). May 1987.

Eisler, R. 2000. *Handbook of Chemical Risk Assessment: Health Hazards to Humans, Plants, and Animals. Volume 2 – Organics*. Lewis Publishers, Boca Raton, FL. ISBN 1-56670-506-1.

Environment Canada. 1998. *Biological Test Method: Reference Method for Determining Acute Lethality of Sediments to Marine or Estuarine Amphipods*. Environmental Protection Series, Report EPS 1/RM/35, December 1998. Environment Canada, Method Development and Application Centre, Ottawa, ON.

EPA (U.S. Environmental Protection Agency). 1989. *Exposure Factors Handbook*. U.S. Environmental Protection Agency, National Center for Environmental Assessment, Washington, DC. August 1989. EPA/600/8-89/043.

EPA (U.S. Environmental Protection Agency). 1991. *Technical Support Document for Water Quality-Based Toxics Control*. EPA 505/2-90-001. Washington, D.C.

EPA (U.S. Environmental Protection Agency). 1992. *Supplemental Guidance to RAGS: Calculating the Concentration Term*. Office of Solid Waste and Emergency Response, U.S. Environmental Protection Agency, Washington, DC. Publication 9285.7-081. http://www.deq.state.or.us/wmc/tank/documents/epa-ucls.pdf (99-0003)

EPA (United Stated Environmental Protection Agency). 1993. *Wildlife Exposures Handbook, Volumes I and II*. Office of Research and Development, Washington, DC. EPA EPA/600/R-93/187a, EPA/600/R-93/187b. (99-0040).

EPA (U.S. Environmental Protection Agency). 1997. *Ecological Risk Assessment Guidance for Superfund: Process for Designing and Conducting Ecological Risk Assessments, Interim Final*. EPA 540-R-97-006. 5 June 1997.

EPA (U.S. Environmental Protection Agency). 1998. *Guidelines for Ecological Risk Assessment*. Risk Assessment Forum, Washington, D.C. EPA/630/R-95/002F.

EPA (U.S. Environmental Protection Agency). 1999. Ecological Risk Assessment and Risk Management Principles for Superfund Sites. Memorandum from Stephen D. Lutwig, Director, Office of Emergency and Remedial Response, to Superfund National Policy Managers, Regions 1-10. OSWER Directive 9285.7-28 P. 7 October 1999.

EPA (U.S. Environmental Protection Agency). 2000a. *Methods for Measuring the Toxicity and Bioaccumulation of Sediment-Associated Contaminants with Freshwater Invertebrates.* 2nd edition. U.S. Environmental Protection Agency, Office of Research and Development, Duluth, MN. EPA/600/R-99/064. March 2000. 192 pp.

EPA (U.S. Environmental Protection Agency) 2000b. *Stressor Identification Guidance Document*. USEPA, Washington, DC.

EPA (U.S. Environmental Protection Agency). 2001. *Risk Assessment Guidance for Superfund (RAGS), Volume III - Part A: Process for Conducting Probabilistic Risk Assessment*. Office of Emergency and Remedial Response, U.S. Environmental Protection Agency, Washington, DC. EPA 540-R-02-002. OSWER 9285.7-45. December 2001. Available on-line at the EPA website http://www.epa.gov/superfund/programs/risk/rags3a/index.htm

EPA (U.S. Environmental Protection Agency). 2002. *Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms*. Fifth edition. EPA-821-R-02-012. October 2002.

EPA (U.S. Environmental Protection Agency). 2002. *Calcasieu Estuary Remedial Investigation/Feasibility Study (RI/FS): Baseline Ecological Risk Assessment (BERA)*. Prepared for CDM Federal Programs Corporation under contract to U.S. EPA, Region 6 by MacDonald Environmental Sciences Ltd., United States Geological Survey and The Cadmus Group, Inc. Contract No. 68-W5-0022, Document Control No. 3282-941-RTZ-RISKZ-14858.

EPA (U.S. Environmental Protection Agency). 2004. *Terms of the Environment*. Revised Edition. EPA Document Order Number 175B97001. Last updated on Friday, January 30th, 2004. http://www.epa.gov/OCEPAterms/

EPA/USACE (U.S. Environmental Protection Agency and U.S. Army Corps of Engineers). 1998. E*valuation of Dredged Material Proposed for Discharge in Waters of the U.S. - Testing Manual. Inland Testing Manual.* U.S. Environmental Protection Agency, Office of Water, Washington, DC and U.S. Army Corps of Engineers, Washington, DC. EPA-823-B-98-004. February 1998.

EVS (EVS Environment Consultants Ltd). 2003. *Assessment of In Situ Stressors and Sediment Toxicity in the Lower Housatonic River*. Final report, adapted from a study by G.A. Burton, 2001. Institute for Environmental Quality. Wright State University, Dayton, OH.

FEL (Fort Environmental Laboratories, Inc.) 2003a. *Final Report-Frog Reproduction and Development Study: 2000 Rana sylvatica Vernal Pool Study*. Prepared for Weston Solutions, Inc. June 2003.

FEL (Fort Environmental Laboratories, Inc.) 2003b. *Final Report-Frog Reproduction and Development Study: 2000 Rana pipiens Reproduction and Development Study*. Prepared for Weston Solutions, Inc. June 2003.

Field, L.J. 1996. *Use of Tissue Residue Data in Exposure and Effects Assessments for Aquatic Organisms*. National Oceanic and Atmospheric Administration, Hazardous Materials Response and Assessment Division, Coastal Resources Coordination Branch, Seattle, Washington. Presented at the National Sediment Bioaccumulation Conference, September 11-13, 1996, Bethesda, Maryland. Pages 2-21 through 2-24. Website location: http://www.epa.gov/waterscience/cs/field.pdf

Fox, G. A. 1991. Practical causal inference for ecoepidemiologists. *Journal of Toxicology and Environmental Health*. 33: 359-373.

Gilbert, M., R. Leclair, Jr., and R. Fortin. 1994. Reproduction of the northern leopard frog (*Rana pipiens*) in floodplain habitat in the Richelieu River, P. Quebec, Canada. *Journal of Herpetology* 28 (4):465-470.

Gill, J.L. 1990. Repeated measurements on animals. Why not assume autocorrelation? *Animal Breed. Gen*. 107: 89-95 pp.

Grapentine, L., J. Anderson, D. Boyd, G.A. Burton, C. DeBarros, G. Johnson, C. Marvin, D. Milani, S. Painter, T. Pascoe, T. Reynoldson, L. Richman, K. Solomon, and P.M. Chapman. 2002. A decision making framework for sediment assessment developed for the Great Lakes. *Human and Ecological Risk Assessment*. 8(7):1641-1655.

Green, R. H. 1979. *Sampling Design and Statistical Methods for Environmental Biologists*. John Wiley & Sons, NY. 257pp.

Haimes, Y.Y., T. Barry, and J.H. Lambert. 1994. When and how can you specify a probability distribution when you don't know much? *Risk Analysi*s. 14:661-706 pp.

Hattis, D. and D.E. Burmaster. 1994. Assessment of variability and uncertainty distributions for practical risk analyses. *Risk Assessment*. 14:713-730.

Heaton, S.N., S.J. Bursian, J.P. Giesy, D.E. Tillitt, J.A. Render, P.D. Jones, D.A. Verbrugge, T.J. Kubiak, and R.J. Aulerich. 1995. Dietary exposure of mink to carp from Saginaw Bay, Michigan. 1. Effects on reproduction and survival, and the potential risks to wild mink populations. *Archives of Environmental Contamination and Toxicology* 28:334-343.

Heath, A.G., 1990. Summary and perspectives. In: *Biological Indicators of Stress*, S.M. Adams (Ed.), American Fisheries Society Symposium 8:183-191.

Hine, R.L., B.L. Les, and B.F. Hellmich. 1981. *Leopard Frog Populations and Mortality in Wisconsin, 1974-76*. Wis. Dep. Nat. Resour., Tech. Bull. No. 122. As cited in Gilbert et al. 1994.

Hoeckstra, J.A. and P.H. Van Ewijk. 1993. Alternatives for the no-observed-effect level. *Environ. Toxicol. Chem*. 12: 187-194.

Ingersoll, C.G., P.S. Haverland, E.L. Brunson, T.J. Canfield, F.J Dwyer, C.E. Henke, and N.E. Kemble. 1996. Calculation and evaluation of sediment effect concentrations for the amphipod Hyalella azteca and the midge Chironomus riparius. *J. Great Lakes Re*s. 22:602-623.

Ingersoll, C.G., E.L. Brunson, F.J. Dwyer, D.K. Hardesty, and N.E. Kemble. 1998. Use of sublethal endpoints in sediment toxicity tests with the amphipod Hyalella azteca. *Environ. Toxicol. Chem.* 17:1508-1523.

Jarvinen, A.W. and G.T. Ankley. 1999. *Linkage of Effects to Tissue Residues: Development of a Comprehensive Database for Aquatic Organisms Exposed to Inorganic and Organic Chemicals*. SETAC Press, Pensacola, FL. 364 pp.

Johnston, R.K., W. R. Munns Jr., P.L. Tyler, P. Marajh-Whittemore, K. Finkelstein, K. Munney, F. T. Short, A. Melville, and S.P. Hahn. 2002. Weighing the evidence of ecological risk from chemical contamination in the estuarine environment adjacent to the Portsmouth Naval Shipyard, Kittery, Maine, USA. *Environ. Tox. Chem.* 21(1):182-194.

Karickhoff, S.W. 1984. Organic pollutant sorption in aquatic systems. *J. Hydraulic Div.* ASCE. 110(6):707-735.

Lee, G.F. and A. Jones-Lee. 2002. *Appropriate Use of Chemical Information in a Best Professional Judgment Triad Weight of Evidence Evaluation of Sediment Quality*. Report of G. Fred Lee & Associates, El Macero, CA. Published in part in the 2000 Proceedings of Sediment Quality Assessment (SQA5), Aquatic Ecosystem Health and Management Society, Chicago, IL.

Legendre, P. and L. Legendre. 1998. *Numerical Ecology*. Second edition. Elsevier, Amsterdam, The Netherlands.

MacDonald, D.D. 1994. *Approach to the Assessment of Sediment Quality in Florida Coastal Waters. Volume 1 – Development and Evaluation of Sediment Quality Assessment Guidelines*. Prepared for Florida Department of Environmental Protection, Office of Water Policy. November 1994.

Malins, D.C., M.M. Krahn, M.S. Myers, L.D. Rhodes, D.W. Brown, C.A. Krone, B.B. McCain, and S.-L. Chan. 1985. Toxic chemicals in sediments and biota from a creosote-polluted harbor: Relationships with hepatic neoplasms and other hepatic lesions in English sole. *Carcinogenesis*. 10: 1463-1469.

Malins, D.C., B.B. McCain, M.S. Myers, D.W. Brown, M.M. Krahn, W.T. Roubal, M.H Schiewe, J.T. Landahl, and S.-L. Chan. 1987. Field and laboratory studies of the etiology of liver neoplasms in marine fish from Puget Sound. *Environmental Health Perspectives* 71: 5-16.

Mapstone, Bruce D. 1995. Scalable Decision Rules for Environmental Impact Studies: Effect Size, Type I, and Type II Errors, *Ecological Applications* 5(2): 401-410.

Massachusetts Division of Fisheries and Wildlife. 2003. 321 CMR 10.00 and 11.00. Massachusetts Endangered Species Act Regulations. Boston, MA.

Massachusetts Office of Environmental Affairs. 1999. 310 CMR 40.00. Massachusetts Contingency Plan (MCP). Boston, MA.

May, R.M. 1975. Patterns of species abundance and diversity. In M.L. Cody and J.M. Diamond (eds.). *Ecology and Evolution of Communities*. pp. 81-120; Belknap Press, Cambridge, MA.

McKim, J.M., G.F. Olson, G.W. Holcombe, and E.P. Hunt. 1976. Long-term effects of methylmercuric chloride on three generations of brook trout (Salveninus fontinalis): toxicity, accumulation, distribution, and elimination. *J. Fish. Res. Board Can.* 33: 2726-2739.

Menzie, C., M.H. Henning, J. Cura, K. Finkelstein, J. Gentile, J. Maughan, D. Mitchell, S. Petron, B. Potocki, S. Svirsky, and P. Tyler. 1996. Special report of the Massachusetts weight-of-evidence workgroup: A weight-of-evidence approach for evaluating ecological risk. *Human and Ecological Risk Assessment*. 2(2):277-304.

Moriarty, F. 1999. *Ecotoxicology*, 3rd Edition. Academic Press, NY.

Moore, D.R.J. and P.-Y. Caux. 1997. Estimating low toxic effects. Environ. Toxicol. Chem. 16: 794-801.

Nichols, J.W., C.P. Larsen, M.E. McDonald, G.J. Niemi, and G.T. Ankley. 1995. Bioenergetics-based model for accumulation of polychlorinated biphenyls by nesting tree swallows, Tachycineta bicolor. *Environmental Science and Technolog*y. 29:604-612.

Niimi, A.J. 1983. Biological and toxicological effects of environmental contaminants in fish and their eggs. *Can. J. Fish. Aquat. Sci.* 40: 306-312.

NRC (National Research Council). 2001. *A Risk-Management Strategy for PCB-Contaminated Sediments.* Committee on Remediation of PCB-Contaminated Sediments. National Academy Press, Washington, DC. 405 p.

Oris, J.T. and A.J. Bailer. 1993. Statistical analysis of the Ceriodaphnia toxicity test: Sample size determination for reproductive effects. *Environ. Toxicol. Chem*. 12: 85-90.

Ott, WR. 1995. *Environmental Statistics and Data Analysis*. Lewis Publishers, Boca Raton, FL.

Ouellet, M. 2000. Chapter 11, Amphibian Deformities: Current State of Knowledge. In *Ecotoxicology of Amphibians and Reptiles*. Society of Environmental Toxicology and Chemistry (SETAC), Pensacola, FL. 904 p.

Papoulias, D.M., S.A. Villalobos, J. Meadows, D. B. Noltie, J.P. Giesy, and D.E. Tillitt. 2003. In ovo exposure to o,p´-DDE affects sexual development but not sexual differentiation in Japanese medaka (Oryzias latipes). *Environmental Health Perspectives*. 111:29-32.

Parkhurst, D.F. 1990. Statistical hypothesis tests and statistical power in pure and applied science. In: *Acting Under Uncertainty: Multidisciplinary Conceptions*, pp. 181-201. Kluwer Academic Publishers, Boston, MA.

Pittinger C. A., J.G. Turnley, and P.M. Mehrle. 2001. A multistakeholder ecological risk management framework. In: *Risk Management: Ecological Risk-Based Decision Making*. (eds R. G. Stahl, R. Bachman, A. L. Barton, J. R. Clark, P. L. deFur, S. J. Ells, C. A. Pittinger, M. W. Slimak, and R. S. Wentsel). SETAC Press, Pensacola, FL. pp. 2132.

Platt, J.R. 1964. Strong inference. *Science* 146, 347-353.

Popper, K.R. 1968. *The Logic of Scientific Discover*y. New York, Harper and Row.

Posthuma, L., G.W. Suter, II, and T.P. Traas. 2002. *Species Sensitivity Distributions in Ecotoxicology*. Lewis Publishers, Boca Raton, FL. (Chapt. 9, pp. 155 – 193).

PSDDA (Puget Sound Dredged Disposal Analysis). 2000. *Dredged Material Evaluation and Disposal Procedures – A User's Manual for the Puget Sound Dredged Disposal Analysis (PSDDA) Program*. U.S. Army Corps of Engineers, Seattle District; U.S. Environmental Protection Agency, Region X, Seattle; Washington Department of Natural Resources; and Washington Department of Ecology, Olympia, WA. February 2000.

Quinn, G.P. and M.J. Keough. 2002. *Experimental Design and Data Analysis for Biologists*. Cambridge University Press, Cambridge, UK. (Chapter 14, pp. 380 – 400).

Resetarits, W.J. 2002. *Experimental Analysis of the Context-Dependent Effects of Early Life-Stage PCB Exposure on Rana sylvatica*. Report prepared for General Electric Company.

Rodríguez-Burgos, A. 2003. Detection in chick embryo of fetoproteins not recognized by the dam's immune system and of soluble alloantigens. Presumptive teratogenic and abortogenic capacity of their specific IgY. *BMC Immunology*. 4:6

Rosenshield, M.L., M.B. Jofre, and W.H. Karasay. 1999. Effects of polychlorinated biphenyl 126 on green frog (*Rana clamitans*) and leopard frog (*Rana pipiens*) hatching success, development, and metamorphosis. *Environmental Toxicology and Chemistry*. 18(11):2478-2486.

Schulz, T.W. and S. Griffin. 1999. Estimating risk assessment exposure point concentrations when the data are not normal or lognormal. *Risk Analysis*. 19:577-584 pp.

Seiler, F.A. and J.L. Alvarez. 1996. On the selection of distributions for stochastic variables. *Risk Analysis* 16: 5-18.

Sibley, P.K., P.D. Monson, and G.T. Ankley. 1997. The effect of gut contents on dry weight estimates of *Chironomus tentans* larvae: Implications for interpreting toxicity in freshwater sediment toxicity tests. *Environmental Toxicology and Chemistry*. 16:1721-1726.

Singh, A.K., A. Singh, and M. Engelhardt. 1997. *The Lognormal Distribution in Environmental Applications*. EPA/600/R-97/006.

Sokal, R.R. and F.J. Rohlf. 1969. *Biometr*y. W.H. Freeman and Company, New York.

Sokal, R.R. and F.J. Rohlf. 1981. *Biometry* (2^{nd} Edition). W.H. Freeman and Company, New York.

Stahl R. G., R. Bachman, A.L. Barton, J.R. Clark, P.L. deFur, S.J. Ells, C.A. Pittinger, M.W. Slimak, and R.S. Wentsel. 2001. *Risk Management: Ecological Risk-Based Decision Making*. SETAC Press, Pensacola, FL.

Suter II, G.W. 1996. Abuse of Hypothesis Testing Statistics in Ecological Risk Assessment. *Human and Ecological Risk Assessment* 2(2):331-347.

Suter II, G.W., R.A. Efroymson, B.E. Sample, and D.S. Jones. 2000. *Ecological Risk Assessment for Contaminated Site*s. Lewis Publishers. Boca Raton, FL.

Tabachnick, B.G. and L.S. Fidell. 1996. *Using Multivariate Statistics*, 3rd Ed. Harper Collins College Publishers, NY. (Chapter 6, pp. 195 – 238)

Thorp, J.H. and A.P. Covich. 1991. *Ecology and Classification of North American Freshwater Invertebrates*. Academic Press, San Diego.

Tillitt, D.E., S.R. Birke, and D.K. Nicks. 2003. Hepatic O-dealkylase activities from induction of cytochrome P-450 in mink and their offspring fed diets containing contaminated fish from the Housatonic River USGS, Columbia Environmental Research Center, Biochemistry and Physiology Branch, Final Laboratory Report FY 2003-30-4, 2003. 16 p.

Van den Brink, P. J. and C.J.F. Ter Braak. 1999. Principal response curves: Analysis of time-dependent multivariate responses of biological community to stress. *Environ. Tox. and Chem*., 18, 138.

Van Wijngaarden, R. P. A., P.J. Van den Brink, J.H. Oude Voshaar, and P. Leeuwangh. 1995. Ordination techniques for analyzing response of biological communities to toxic stress in experimental ecosystems, *Ecotoxicology*, 4, 61.

Varanasi, U., J.E. Stein, and M. Nishimoto. 1989. Biotransformation and Disposition of Polycyclic Aromatic Hydrocarbons (PAHs) in Fish. In U. Varanasi, ed., *Metabolism of Polycyclic Aromatic Hydrocarbons in the Aquatic Environmen*t, CRC Press, Boca Raton, FL, pp. 94-149.

Varanasi, U., J.E. Stein, L.L. Johnson, T.K. Collier, E. Casillas, and M.S. Myers. 1992. Evaluation of bioindicators of contaminant exposure and effects in coastal ecosystems. In D.H. McKenzie, D.E. Hyatt, and V.J. McDonald, eds., *Ecological Indicators*, Vol. 1. Proceedings of an international symposium, Fort Lauderdale, Florida, October 16-19, 1990.

Wiener, J.G. and D.J. Spry. 1996. Toxicological significance of mercury in freshwater fish. pp. 297-340. In: W.N. Beyer, G.H. Heinz, and A.W. Redmon-Norwood (Eds.). *Environmental Contaminants in Wildlife: Interpreting Tissue Concentrations*. Society of Environmental Toxicology and Chemistry, Lewis Publishers.

WESTON (Roy F. Weston, Inc.). 2000. *Supplemental Investigation Work Plan for the Lower Housatonic River*. Prepared for U.S. Army Corps of Engineers and U.S. Environmental Protection Agency. 22 February 2000. DCN: GEP2-020900-AAME. (07-0078).

WESTON (Weston Solutions, Inc.). 2003. *EPA's Final Response to the ERA Peer Review Panel Initial Questions and Comments.* Prepared for U.S. Environmental Protection Agency and U.S. Army Corps of Engineers. 11 December 2003. DCN: GE-121103-ABYN.

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