Deriving Sediment Interstitial Water Remediation Goals (IWRGs) at Superfund Sites for the Protection of Benthic Organisms from Direct Toxicity

Peer Review Charge Questions

Background Information:

Over the past two decades, methods for measuring the concentrations of bioavailable chemical in sediments have been developed. Research has shown that the bioavailable chemical in sediment and freely dissolved chemical in the sediment interstitial water are practically equivalent. This document provides a methodology for deriving interstitial water remediation goals (IWRGs) based upon the bioavailable/freely dissolved chemical in the sediment interstitial water for the protection of benthic organisms from direct toxicity. Remediation goals are derived on a sediment interstitial water basis (µg/L) and subsequently, are converted to a bulk sediment basis (µg/kg dry weight) using site-specific sediment/water partition coefficients. Additionally, this document contains guidance on how to compare and evaluate results from sediment toxicity tests to concentrations of chemical in the sediment interstitial water. When these two results are consistent with each other, one can be reasonably assured that the causes of toxicity to benthic organisms in the sediment have been correctly identified and that the developed IWRGs for the toxicants will be protective of the benthic organisms at the site. The consistency evaluation is an important step in developing defensible IWRGs.

Charge Questions:

As you read through the sections of this document that you have been asked to review, please provide written responses to the best of your ability to the following questions. Additional comments and recommendations for improving this document and associated methodology are also welcome:

(1) Is the document written in a style that will be accessible for users with a range of educational and technical backgrounds?

The document is generally well written and will be accessible to a wide range of users that are familiar with Superfund site assessments. There are several instances where defining or clarifying terminology or revision text might be helpful as noted in specific comments and proposed text changes included in the attached document.

(2) Is the described methodology sufficiently clear to be performed by Superfund remediation project managers, risk assessors, and consultants for Superfund sites? If not, please provide suggestions on how clarity can be improved.

I think the general 4 step outline describing the proposed methodology is clear. A key practical challenge is step 4 and I think that the authors need to make the point that this step may not be required particularly in light of recent advances in sediment remedies that focus on in-situ amendments where targeting reduction in Cfree is the remedial objective (not mass based sediment concentrations). Further, the efficacy of the remedial action can be confirmed using passive sampling as a monitoring tool. This strategy appears to be overlooked in the present report and should be discussed in section 4 before proceeding to describing approaches used for step 4 which may add significantly uncertainty that could undermine the advantages of applying IWRGs for improved sediment remedial decision-making.

I feel table Table 3-1 could be streamlined by presenting a single recommended IWRG (or two values if separate freshwater and marine values) that is intended to provide a chronic protection level rather than presenting multiple values (i.e. SCVs, FCVs, ESBs). This will avoid confusion and ensure more consistent application of the contaminant-specific IWRGs that are presented.

I also suggest that the authors consider preparing a table of IWRGs corresponding to key sediment test organisms/endpoints for the NOAA 34 PAHs that can then be used for calculating ∑TUs that can be compared to observed toxicity data. This will facilitate consistency by users of the this guidance document in evaluating relationship between site-specific chemistry and toxicity data as described in section 5-3. An alternative would be to provide a simple spreadsheet tool that users could apply for this purpose.

(3) Is the document missing any important concepts, sections, definitions, and/or text that should be provided in order to make the methodology truly implementable?

A key deficiency is a discussion of the two key formats for passive sampling (ex-situ vs in-situ). This issue is briefly mentioned in section 5 but given the importance of sampling format in practical implementation of this technology in Superfund site assessments this deserves more discussion. Which format to apply should consider both objectives of the study relative to the pros/cons of each sampling approach. If the objective is to compare passive sampling results to lab toxicity tests, ex-situ measurements are preferred since they are cheaper and can be performed under more controlled conditions that facilitate equilibrium and translation in reliable Cfree measurements. If instead the objective is to compare passive sampling results to observed impacts on field macroinvertebrate communities (or calibration of a site-specific bioaccumulation model) then in-situ measurements may be preferable since reliable estimates of actual Cfree concentrations under field conditions are more essential to the study objective. As far as I know, limited information in comparing ex-situ vs in-situ site data are available so if reliable estimates of field measurements are needed then an initial study assessing concordance between approaches may be warranted in guiding the definitive study design.

In section 2.1 the authors state “measurements from compromised sampler must not be used.” However, little practical guidance is provided to determine when to judge measurements as comprised. It would be helpful to provide some general criteria: e.g. highly variable results between

replicates; predicted Cfree concentrations exceeding solubility; chromatograms that are characteristic of oil present in the sediment.

Section 4.2 should also mention that if an evaluation of OC normalization indicates variability in site-specific partitioning of a contaminant is not reduced when compared to dry weight normalization than OC normalization may add little value in the translation step. Further, the potential use of probabilistic methods should also be acknowledged for evaluating the uncertainty site-specific sediment-water partition coefficients if translation to sediment concentrations are required (currently only a deterministic approach is discussed).

Section 4.4 should include an option to include the IWRG as the basis for the remedial decision (not include a translation step that allows the significant uncertainties discussed to be circumvented). If a translation step is included then additional guidance to evaluate key assumption that porewater composition of dissolved phase constituents is would be helpful (e.g. prepare bar charts to visually show relative composition of porewaterPAHs at different total concentrations)

Section 5 indicates highlights three types of replicates should be considered but this specifically relates to application of an ex-situ sampling format. Replication for in-situ sampling should also be considered. It is also suggested that the authors may wish to contact Dr. Chiel Jonker who has recently completed a rather extensive inter-laboratory comparison evaluation of ex-situ passive sampling measurements for sediment PCBs and PAHs. The results of this exercise may provide insights on the expected magnitude of variances in Cfree estimates observed between labs, locations, batches within a location and replicate passive sampler measurements.

(4) Are the illustrative examples for determining IWRGs complete enough to demonstrate how the IWRGs are derived?

The authors have attempted to include some useful illustrative examples. A more detailed case study that describes the step by step application of this approach to a specific site and highlights the significant impact of this approach over the default EqP paradigm in deriving sediment remedial goals would be welcomed. However, this may be difficult given publically available site data may not yet be available for this purpose.

(5) Is the methodology for deriving interstitial water remediation goals scientifically defensible?

The methodology for establishing IWRGs is based on earlier peer review publications so is considered scientifically defensible (e.g. Burgess et al. 2013). For PAHs, more recent work by Redman et al. can 2014 be cited to further support application of the TLM for chronic protection of benthic organisms. The document offers limited new guidance for establishing IWRGs for additional contaminants of concern other than generating water-only toxicity tests for establishing a species-sensitivity distribution which is costly and may be impractical. However, recent advances in extending the target lipid model using polyparameter linear free energy relationships has a much

wider chemical domain and could be mentioned as a promising future modeling tool for potentially deriving IWRGs for emerging contaminants of concern for which limited toxicity data are available.

One recommendation provided in section 2.1.1 of the report that I believe lacks sufficient technical justification was that the ASTM/EPA SPME method is the best approach for analysis of sediments samples that may be confounded by NAPL contamination. The authors provide little technical basis to show that this technique would not yield measurements that are similarly "compromised". Unless further data can be provided to support this position, it is recommended that the authors simply present this an alternate method that can be considered. The principle advantage of this method is that a standardized test methodology is available. However, this method is not directly comparable to equilibrium sampling and to my knowledge few labs other than Hawthorne perform this method.

(6) In implementing the methodology, site-specific K_{OC}s are used to convert the IWRGs on concentration basis in sediment interstitial water (µg/L) to concentrations in bulk sediment (µg/kg dry weight). Is the discussion of the K_{OC}s adequate? Is the discussion of the conversion from concentrations in interstitial water to bulk sediment adequate? Is the discussion of which K_{OC}s should be used in the conversions adequate?

I suggest that the authors add two elements to this discussion: evaluating the need for OC normalization and the potential use of probabilistic methods for evaluating the uncertainty site-specific sediment-water partition coefficients in translation to sediment concentrations when this step is needed (see comment 3 above).

(7) Passive sampling can be performed on any number of samples from a site; for example, on all samples where contaminants are measured in bulk sediment, on only the surface sediments, on the top and bottom of sediments cores, on the top and at the dredge depth of the sediments cores, on surface sediment and based of BAZ (biological active zone), or some other arrangement. Currently, the methodology allows flexibility (makes no recommendation) on which samples are measured using the passive sampling technique and how those data are used in the conversion from interstitial water IWRGs to bulk sediment IWRGs. The extremes in this process are a) perform one passive sampling measurement and assume all sediments are the same across the location of interest (horizontally and with depth) or b) perform passive sampling on all samples and develop 3-D contour plots with depth based upon concentrations in the interstitial water. Should the methodology make a recommendation on this issue? If so, provide your recommendation.

I agree the guidance should not be too prescriptive given limited practical experience is available in applying this approach to date. It may be helpful to emphasize with some examples that the scope of applying this approach will vary based on study objective. For example, a screening site risk assessment using conventional total sediment concentrations that are organic carbon normalized may indicate that based on EqP assumptions only a very limited spatial extent of sediment appears to pose a potential concern. This area could then logically be the focus of a targeted follow-up study where passive sampling and complimentary effects data (field surveys of benthic health or toxicity

tests) are collected. In contrast, if potential risks appear widespread based on conventional characterization of sediment contamination more extensive use of passive sampling may be warranted that includes not only samples from the site but also reference stations so that the comparative bioavailability of contaminants in site sediment can be compared to EqP assumptions and potentially differentiated from reference conditions. Further, at sites where there is a large variation in the magnitude of a sediment contaminant concentrations a key study objective may be to define how bioavailability changes as a function of total sediment contamination since this information will be critical for remedial design. These specifics of the study design will also depend on a variety of practical considerations including cost and time trade-offs and receptivity of the EPA region or state and potentially responsible parties to generate and apply these data in decision-making.

(8) Section 5 provides information on comparing toxicity test results and developed IWRGs. Is this section sufficiently clear for the non-experts in toxicity testing and/or passive sampling?

Please see earlier response to comments 2 and 3 that provide some suggestions for improving section 5. In addition, some readers may not understand Figure 5-2 which does not follow the format of Figure 5-1 in which the x-axis is expressed in terms of toxic units. It would be clearer if it was possible to depict Figure 5-2 as a two panel plot where in the first panel survival vs TUs based on FCV were plotted and on the second panel survival vs TUs based on hyallella acute toxicity was plotted. This would allow you to then make point that later plot is more appropriate for comparison to the empirical toxicity data as indicated in the position of the concentration-response relationship since the TU used reflects the sensitivity of the organism tested. If this is not possible, I suggest adding text to point out that blue dotted line in the current version of Figure 5-2 corresponds to acute critical body burden for *Hyalella* and shows consistency with the position in the observed concentration response.

Several additional comments on specific sections of the report and suggested editoral text changes of the accompanying marked up version of the document is also provided for consideration by the authors.

Please provide your written comments to Virginia Houk (Houk.virginia@epa.gov) no later than July 15, 2016.

If you have any questions concerning the draft guidance or the charge, please do not hesitate to contact me at 919-541-2815. We sincerely thank you for your input to our peer review process.

Virginia S. Houk
Peer Review Coordinator / Designated Federal Officer
USEPA/NHEERL
Maildrop B305-02
Research Triangle Park, NC 27711
T: 919.541.2815

houk.virginia@epa.gov