Comments from Philip A. Clifford, Ph.D.

QUESTIONS

Framework

1) Question:

Is the EPA Region2/CENAN Framework for evaluating bioaccumulation results scientifically appropriate for determining the suitability of dredged material as Remediation Material? If not, describe deficiencies. *(Please see Region2lCENAN joint evaluation memorandum, Figure 1)*

1) Response:

This framework is one of the more thorough attempts to directly evaluate body burden analytical data that I have seen. As such, I believe that, if direct interpretation of such data is the assessment method of choice, it can be accepted with few reservations. Scrutiny of the data set used suggests to me that the hypothetical sediment presented would indeed be appropriately identified as category I and, as such, the screening methodology seems to have worked. However, I have some discomfort regarding the inherently large uncertainties surrounding direct toxicological interpretation of body burden data (discussed in more detail in specific responses below). As there are several ways in which body burden data can be used in ecological risk assessment I recommend consideration of some of the alternatives.

One of the principal alternatives is to use the data as input to a dose-based food-web model rather than attempting to toxicologically evaluate the data themselves. Given that concentrations of contaminants in sediment, water, and lower trophic level organism tissues are either known or can be easily predicted using conventional partitioning methodologies, it is a relatively straight forward task to estimate the doses of contaminants that receptors virtually anywhere in the food-web are exposed to. For example, using widely accepted models like those of Gobas, Thomann, Connelly, Parkerton, et al., the existing data are sufficient to predict concentrations of contaminants in fish and risk to piscivorous birds (e.g., pelican) could be estimated by comparing calculated dietary intake to good benchmark studies. Risks at other trophic levels can be evaluated in a similar manner. I also believe that evaluation of such upper trophic-level receptors would be more appropriate than using humans as terminal receptors for ecological assessment. Methods and modeling parameters for estimating dietary intake to ecological receptors are presented in USEPA (1993) and elsewhere and there are numerous peer-reviewed articles that present good dose-based toxicological benchmark data for comparison. I believe that conduct of such an analysis could be either a very robust compliment to the existing framework or could be used to replace portions of it.

Whether dose-based evaluations are considered or not, one of the principal things I recommend adding to the framework and draft memorandum for the record is a specific section presenting a thorough treatment of uncertainties. While it is clearly not possible to quantitatively address all of the uncertainties associated with the comparisons presented, some qualitative treatment should be incorporated in the document for the information of the decision makers. For example,

regional matrix values were derived by calculating tissue concentrations from grand means of concentration data and a conservative (lowest) BCF. While I do not argue here with the validity of that approach, uncertainties associated with the derivation method are critical to understanding the meaning of exceeding such a value. In this case, treatment of concentration data is non-conservative because a mean has been used but the overall calculation is conservative because of the BCF selected. The point in this particular case is that populations of organisms should be protected (use of the grand mean) but not all individuals of those populations will necessarily be protected. It is very important to understand these kinds of issues when interpreting the meaning of the comparisons presented.

Brief discussions of the uncertainties (and degrees of conservatism) associated with the various comparison criteria could be presented prior to section VI and used as an integral part of drawing the overall conclusions. This would be more in keeping with the "weight of evidence" (or lines of evidence) approach generally preferred for ecological risk assessment. I do not believe that there is any need to waste time trying to educate readers of the document regarding the inherent uncertainties associated with things like selection of "=0.05 for statistical testing as that sort of information is widely available in the general literature, but a more thorough qualitative treatment of other uncertainties could improve overall confidence in the conclusions reached.

2) Which of the risk-based values derived constitute "true" conservative estimates of risk levels (i.e., exceeding the value should be interpreted as sufficient cause to conclude that significant undesirable effects may result through bioaccumulation)? Which of the risk-based values derived constitute conservative screening values (i.e., test tissue concentrations below the value can confidently be interpreted to pose no risk of significant undesirable effects and exceeding should be further evaluated before the probability of significant undesirable effects can be assessed)? How can the "true" risk levels be calculated for those compounds which you believe only to have screening values? How should test concentrations be compared to risk-based levels to determine whether they are exceeded?

2) Response:

Because I believe that all of the comparison criteria (a^1, b^1, b^2, b^3, c) have "risk-based" components, I will not restrict my comments to those values presented as "risk-based" in section V.C.2.c. Also, note that by definition, a conservative estimate of risk is not a "true" level of risk, but one that is biased toward conservatism.

As I'm not entirely comfortable that any body burden evaluations really constitute true estimates of risk (significant undesirable effects may not result through bioaccumulation), it may be appropriate to designate exceedance of the screening values presented as "constituting sufficient cause for further evaluation", rather than as "constituting actual risk of significant undesirable effects".

As to methods of comparison of test concentrations to risk-based levels, I have no difficulty with the statistical methods used, although non-parametric methods might be preferable to parametric ones given the likely large departure of environmental conditions from those required for parametric hypothesis testing.

a) Comparison of bioaccumulation test results to reference sediment test results

Comparison of site results to reference area results is probably the least biased of the screening values (closest to a "true" estimate), although sample size could be an issue under some circumstances. As noted in the text, exceedance of reference values is common when those reference values contain a lot of non-detects. This difficulty might be reduced to some degree by replacing the "half-detection limit" method employed with a technique like maximum likelihood estimation to deal more effectively with the censored (below detection limit) tails of the distributions (e.g., Lindgren 1976; Sielken et al. 1993; Clifford et al. 1995; Banton et al. 1996). Also, as I discuss in more detail in other responses herein, I am not entirely convinced that a body burden of a contaminant can necessarily always be related to a toxicological effect. As such, while the comparisons are good for identifying circumstances where site sediments have greater chemical activity than reference sediments, caution must be exercised when interpreting the meaning of the results.

There is also the issue of contaminants potentially present but not analyzed for. This is a vexing problem in most risk assessments and not one we are likely to solve here but, I suggest some qualitative language in an uncertainty section and perhaps a discussion early in the document regarding how the analyte list was generated.

b[1]) Comparison to FDA Action levels

FDA Action levels are not generally derived for protection of the environment as human health and economic concerns are paramount therein. As such, I do not believe that they are really appropriate for the purposes of this sort of an evaluation (regardless of the applicable regulations), although their inclusion does not in any way compromise the process as they constitute a reason to reject a sediment rather than a reason to accept one. I believe that these levels are probably, in general, the least representative of "true" ecological risk estimates of those presented.

b[2]) Comparison to Regional Matrix levels

The regional matrix values were derived by very conservative methods (e.g., selection of the lowest available BCF). As such, they do not represent "true" risk levels, but conservative ones. In addition, because they are calculated values, not empirically measured ones, the associated uncertainty cannot really be evaluated. These values, if we accept the notion that body burdens can be directly related to effects, are probably the most useful as screening values and would be the closest of the values presented to ones that when not exceeded confidently represent category I sediments.

b[3]) Comparison to Regional Dioxin Values

I'm uncertain regarding the degree of conservatism (or accuracy) in the regional dioxin values for ecological concerns. This is principally because the values are derived for protection of human health and humans may not be an appropriate end receptor for evaluation of ecological risk. The solitary study compared to for actual ecological concerns is that of Cook et al. (1993) which presents a value of 50 pptr as a "low risk" concentration for adverse effects on fish while the relationships used to derive the criterion predict a value of 20 pptr in fish tissue at the criterion.

This margin of safety (factor of 2.5) may actually represent a value within the range of chronic toxicity, but this is unknown. I would have to see much more environmental data here and consideration of or more applicable ecological receptors to be comfortable with the uncertainty in the associated values. I believe that it would be appropriate to select an upper trophic level receptor that is more relevant to the site than humans like a large carnivorous fish or a piscivorous bird.

c[1]) Risk-based consideration of bioaccumulation and food-chain transfer potential

Although both high bioaccumulation and high food-web transfer potentials can generally be considered to be undesirable, they are not, in a strict sense, a measure of the environmental risk posed by a compound. According to the studies of many researchers, the rate at which an organism receives it's exposure (it's dose rate) may be of greater importance than the ultimate body burden. PAHs are a good example here as the rate of uptake (dose rate or exposure rate) may be a good predictor of effects while body burdens may remain fairly constant at low values due to metabolism. The notion of "steady state" with such easily metabolized compounds may be illusory at best and misleading at worst.

There are certainly researchers who maintain that body burdens for many contaminants can be confidently related to adverse effects and, although I do not propose to present a data-laden debate on that subject herein, I have much greater confidence in dose estimates than body burden estimates based on the data I have encountered. Also, for many high Kow compounds (e.g., DDTr and dioxins), "steady state" may never actually be achieved by any exposed organism (Pruell et al. 1990 notwithstanding) because the depuration (including transformation, metabolism, and elimination) rate is essentially zero and tissue concentrations never actually reach a maximum (although there may be a functional upper boundary where lethality always occurs). For other compounds like PAHs, bioaccumulation and body burden may be very difficult to interpret due to metabolism. As such, while it is certainly possible to predict bioaccumulation for some compounds, and this can be useful for modeling contaminant movement in a food-web, I am not entirely comfortable that risk should be solely evaluated on this basis.

c[2]) Risk-based comparison to background concentrations

Although simple comparison of concentrations to background is not terribly robust, it is probably one of the more accurate in terms of identifying potential for risk. This statement must, of course, be taken with the cautions presented above regarding bioaccumulation. The greatest difficulty here will be identification of a true "background" location. Note also that for a very clean background sediment, the propensity of non-detects will drive numerous spurious exceedances by project sediments where risks do not actually exist. Nonetheless, I believe that considered comparison of actual data with actual data are the most accurate of the screening techniques presented herein.

c[3]) Risk-based potential for ecological effects

Body burdens are again employed herein and the same reservations I have expressed above are applicable here. My concerns regarding toxicological interpretation of body burdens aside though, the statement that "CBRs are represented as the ratio of the mass of the toxicant per

kilogram (mmole or ug/kg) of organism." for PAHs is not entirely consistent. If we accept the notion that one molecule of a PAH is toxicologically equivalent to any other molecule of a PAH, (which is not without merit and supporting data) then the measurements must be made on a mmole basis, not on a mass basis. The following conversions (using BaP as a standard since the most toxicological information is available for that compound) would be required:

PAH	Molecular Weight	Conversion Factor
Naphthalene	128.16	1.97
Acenaphthylene	152.21	1.66
Acenaphthene	154.21	1.64
Fluorene	166.21	1.52
Phenanthrene	178.22	1.42
Anthracene	178.22	1.42
Fluoranthene	202.26	1.25
Pyrene	202.24	1.25
Benz[a]anthracene	228.28	1.11
Chrysene	228.28	1.11
Benzo[b]fluoranthene	252.32	1.00
Benzo[k]fluoranthene	252.32	1.00
Benzo[a]pyrene	252.32	1.00
Dibenz[a,h]anthracene	228.28	1.11
Indeno[1,2,3-cd]pyrene	276.00	0.91
Benzo[g,h,i]perylene	276.00	0.91

Therefore, the statement that a 400 ppb dose of naphthalene is equivalent to a 400 ppb dose of fluorene is incorrect as the equivalent mass of fluorene would be 519 ppb to achieve the same number of molecules. On that basis, the total dose would be 919 ppb naphthalene toxicological equivalents, not 800 total PAH equivalents.

c[4]) Risk-based consideration of potential effects on human health

As I've indicated above, I have reservations that humans are the most exposed or most at-risk upper level consumers at this site and recommend consideration of a possibly more ecologically relevant receptor such as a piscivorous bird or fish.

Benchmark and Risk Evaluation Values

7. Human Health Risk, Cancer and Noncancer

7A) Question:

Are the risk values suitable for determining the suitability for placement at the HARS as Remediation Material? If there are better alternatives for human risk, specifically what are they?

7A) Response:

As stated above, I believe that there may be other receptors that are more ecologically relevant like upper trophic-level fish (e.g., tuna, jack, etc.) or piscivorous birds (e.g., pelican, osprey, gull, etc.) that would be far more exposed than humans, perhaps far more susceptible to impacts (e.g., pelican egg shell thinning with exposure to DDT), and potentially more relevant. I say this, however, not knowing the area well enough to know which of these receptors might be present, if any.

7B) Question:

Benthic tissue levels for cancer protection were derived using assumptions focused on attaining a cancer protection at the 10-4 risk level. Is this risk appropriate for a determination of ocean placement of Remediation Material? *(Please see Region2lCENAN joint evaluation memorandum, Appendix for Table 1, Page A-4, A-5)*

7B) Response:

Assuming acceptance of humans as a terminal receptor, yes.

7C) Question:

7C-1) Benthic tissue levels for noncancer protection were derived using Reference Dose (RfD) of several organic and inorganic contaminants for the protection of human health. Are these values appropriately and consistently derived?

7C-1) Response:

Assuming acceptance of humans as a terminal receptor, they appear to be.

7C-2) Is the whole body/fillet conversion factor of 1.35 an appropriate factor for all of the contaminants considered if human exposure is assumed to be primarily via consumption of the fillet portion of the fish? *(Please see Region2lCENAN joint evaluation memorandum, Appendix for Table 1, Attachments B and C). If* not, what factors would be appropriate?

7C-2) Response:

Within the limits of my expertise in human health risk assessment, yes.

7C-3) For the lead noncancer value, since there is no RFD for lead the value was derived differently than the other metals. Was the value derived appropriately? *(Please see Reference No. 88)*

7C-3) Response:

Within the limits of my expertise in human health risk assessment, yes, albeit very conservative.

7D) Question:

Are the risk values suitable for predicting the significant undesirable effects due to bioaccumulation? Since the primary route of exposure is through consumption of fish and shellfish, should the variability in potential exposure due to differences in fishing behavior (e.g., target species, seasonal preferences) be incorporated in the risk paradigm?

7D) Response:

As stated above, I believe that there may be receptors other than humans that are far more ecologically relevant. As such, I cannot answer the question as such and recommend consideration of a different receptor with due concern given to it's seasonal variability.

8. Ecological Risk

8A) Question:

Ecological effects benchmarks include the Water Quality Criteria Tissue Level (WQCTL), Critical Body Residue (CBR) associated with narcotic responses, and certain mutagenic/teratogenic effects. Is it valid to use the CBR effect end point for evaluating significant undesirable effect? Are there other ecological end points that should be used to measure ecological risk that are protective of marine benthic and fish life via trophic transfer, particularly for PAHS? If so, identify. With regard to a narcotic effect for chlorinated organic compounds, should an additive approach be considered to include the contribution of chlorinated hydrocarbons against this narcotic (CBR) endpoint.

8A) Response:

As discussed above (question 2 and elsewhere) I'd really like to see a complete food-web risk analysis performed which evaluates risks to pertinent trophic levels on an exposure, rather than body burden, basis. This would involve a site-specific food web including molluscs, annelids, arthropods, fish of several trophic levels, and possibly piscivorous birds (depending on site conditions that I am not familiar with). Much of the ground work for such an analysis has already been done here (tabulation of toxicological endpoints like water quality criteria, compilation of bioaccumulation factors, etc.) and should not require that much more effort. The tissue burden approach inherent in the CBR approach is not without merit but, I believe that a more thorough evaluation of the food-web as a whole on an exposure basis may be worth considering.

As to other chlorinated organic compounds, if the principal toxicological effect on target species is known to be narcotic, additivity can certainly be considered, noting my comments on moles vs. milligrams above.

8B) Question:

Is the Region 2 WQCTL approach (i.e., multiplying the Water Quality Criteria Chronic Value by the Bioconcentration factor) appropriate for determining ecological effects levels of the contaminants for which they were developed? Specifically, are the appropriate BCFs used (for fish, bivalves, etc)? *(Please see Region2lCENAN joint evaluation memorandum, Appendix for Table 1, Page A-1)*

8B) Response:

I have no real problem with the concept that exposure at the CWA CV represents an appropriate toxicological threshold. I do, however, as noted elsewhere herein, have reservations regarding the meaning of body burdens as surrogates for exposures.

While I would like to see BCFs specific to the organisms and contaminants in question used in all cases, this is clearly not possible as the data simply do not exist in all cases. I believe that the values used represent a scientifically defensible attempt to arrive at the most applicable values available.

8C) Question:

BCFs reported for fish were used in the calculations of WQCTLs for organics; is this derived level appropriate for setting benthic tissue ecological effects levels? If the fish tissue levels are used, should adjustments be made to the derived levels to reflect the higher lipid contents of the benthic organisms used in the testing program?

8C) Response:

Notwithstanding my comments on the utility of body burdens, the BCFs used are probably the best available as such values may not be available at all for benthos. As such, I think some discussion of the uncertainties is appropriate but, you cannot replace the values used with better

values when none exist. Addressing the uncertainties should be sufficient. Adjustment of the values for lipid content is appropriate for the more hydrophobic compounds.

8D) Question:

Are the WQCTLs calculated for metals using bivalve BCFs appropriate for setting levels for polychaetes or vice versa?

8D) Response:

Again, you are constrained to using the available information. While I would prefer to see values for polychaetes used for polychaetes, rather than bivalve values, if no such values are available, this is a fairly reasonable approximation. Again, however, I urge addition of an uncertainty section to the report which discusses the limitations of such approaches as regards decision making.

8E) Question:

Are the uncertainty factors applied while deriving ecological effects levels for PAH contaminants appropriate? Does this adequately address the uncertainty around the derived values? Can uncertainty be accounted for using these order of magnitude adjustments? Should they be applied elsewhere to the other risk-based values?

8E) Response:

Numerous attempts are currently on-going (some by myself) to derive uncertainty factors more appropriate than those presented. None currently exist. I believe that the approach used is conservative and cannot at this time offer defensible alternatives. I've commented on the overall uncertainties in the approach repeatedly through this document and, while I believe that the uncertainties are "adequately addressed" by these uncertainty factors, I still think that the document would benefit from some qualitative treatment of the overall uncertainties.

I do not believe that such factors should be applied elsewhere. Qualitative discussion of uncertainties prior to drawing conclusions would be appropriate, however.

8F) Question:

Are the risk values suitable for predicting the significant undesirable effects due to bioaccumulation; are there better alternatives for ecological nonspecific risk?

8F) Response:

Please see responses to 8A and 8B (and others throughout).

8G) Question:

If you believe that these values are over- or under- conservative, what do you believe to be an appropriate way to improve them?

8G) Response:

Please see responses to 8A and 8B (and others throughout).

Calculations

11) Question:

Is the calculation and use of BaP toxicity equivalence an appropriate way to estimate the potential carcinogenicity of PAHS? *(Please see Region2lCENAN joint evaluation memorandum, Appendix for Table 1, Section C.)*

11) Response:

My concerns regarding use of body burdens as surrogates for exposure values aside, the conversion technique itself is, in general, appropriate. See comment response #2 regarding moles vs. milligrams.

13) Question:

Is the assumption of a trophic transfer coefficient of one appropriate for use in evaluating the potential for human health and ecological impacts associated with metals in Remediation Material? Are the trophic transfer factors calculated for organic compounds correct? *(Please see Region2lCENAN joint evaluation memorandum, Appendix for Table 1, Attachment C.)*

13) Response:

My concerns regarding use of body burdens as surrogates for exposure values aside, a trophic transfer coefficient of 1.0 may be low. A review of Suedel et al. (1994) suggests that higher metal-specific values may be more appropriate.

14) Question:

Is the assumption of a fish consumption rate of 6.5 g/day appropriate for use in evaluating the potential for human health impacts associated with metals in Remediation Material? *(Please see Region2lCENAN joint evaluation memorandum, Appendix for Table 1, Page* A-5) Would it be appropriate that the evaluation focus on a higher consumption population?

14) Response:

Given my limited expertise in human health risk assessment and setting aside my concerns regarding use of humans as terminal ecological receptors, I believe the value is appropriate.

General

15) Question:

Is it plausible to replace any other risk assessment assumptions with assumptions specific to the HARS site? *(Please see Region2lCENAN joint evaluation memorandum, Appendix for Table 1, Attachment C and Reference Nos. 88).* Is it appropriate to consider the HARS intended use to be factored into an evaluation of effects at the community or population level?

15) Response:

The only site-specific information that I can think of that could be used to replace some of the default assessment assumptions would be regarding the character of the food-web at the HARS site. For example, while the species selected for evaluation follow guidance and are generally sensitive and appropriate, if it were known that more or less sensitive species were present at the site or that top-predators more appropriate for use in risk assessment than humans were potentially at risk, I would recommend their inclusion in this evaluation. I do not believe that the HARS intended use is appropriate for consideration at the risk assessment stage as extant risk is independent of intended future site use. Consideration of these kinds of factors should take place at the risk management stage, not the risk assessment stage.

17) Question:

Should risks from synergistic effects, from exposure to multiple contaminants, be evaluated using results from tissue analyses? If so, how? If not, why not?

17) Response:

Although this is a laudable goal, I believe that it is generally intractable given the current state-of-the-science. Under conditions where such cumulative effects are both understood and quantifiable for the contaminants in question and the species of interest, such attempts should be made. However, since this will be the exception rather than the rule, it may be best to relegate such evaluations to the uncertainty section that I have recommended should be added to this framework. For the time being, the actual toxicity tests (section V.C.1), as they integrate all of these potential cumulative effects, might be given greater weight in the process and testing could be extended to move from the current, essentially acute tests, to test durations more reflective of truly reflective of chronic exposures. In that manner such effects will be accounted for as inherent components of the actual bioassays. To my way of thinking such empirical information is of greater value to a decision-maker than modeled or interpreted evaluations.

References

Banton, M.I., J.S. Klingensmith, D.E. Barchers, P.A. Clifford, D.F. Ludwig, A.M. Macrander, R.L. Sielken, and C. Valdez-Flores. 1996. An approach for estimating ecological risks from organochlorine pesticides to terrestrial organisms at Rocky Mountain Arsenal. Human and Eco. Risk Assmt. 2(3):499-526.

Clifford, P.A., D.E. Barchers, D.F. Ludwig, R.L. Sielken, J.S. Klingensmith, R.V. Graham, and M.I. Banton. 1995. An approach to quantifying spatial components of exposure for ecological risk assessment. Env. Tox. and Chem., 14(5):895-906.

Lindgren, B.W. 1976. Statistical Theory, Third Edition. Macmillan Publishing Co., Inc., p. 191.

Sielken, R.L., C. Valdez-Flores, P.A. Clifford, D.F. Ludwig, J.S. Klingensmith, M.I. Banton. 1993. A co-located distributions approach for estimating the bioaccumulation factor for ecological risk assessment. Presented to: Society of Toxicology, 1993 annual meeting, New Orleans, Louisiana.

U.S. Environmental Protection Agency (USEPA). 1993. Wildlife exposure factors handbook, volumes I and II. Office of Research and Development. EPA/600/R-931/187a&b.