August 24,1998

Peer Review of USEPA/Army Corps of Engineers :

EPA Region 2/CENAN Framework for Evaluating Dredged Material for Proposed Placement at the HARS

S.M. Bartell

The Cadmus Group, Inc.
136 Mitchell Road
Oak Ridge, TN 37830
Questions:

Framework

2. Risk, by definition, is the conditional probability of some undesired event occurring, along with some statement of its consequences (i.e., human health, ecological). In the absence of perfect information, our data and understanding are conditioned on all the sources of bias and imprecision inherent to the scientific enterprise. Thus, there are no "true"values among the risk-based criteria described in the report. Nevertheless, we might begin to approximate some of these kinds of values with sufficient accuracy and reliability that they can be justifiably used in a decision-making framework (e.g., Figure 1 of the report). For example, the sediment concentrations back-calculated using the WQCTL and the BCF method might provide initial estimates of exposure potentials that would lead to unacceptable risk. This approach might be made more conservative (i.e., pessimistic) by using the lowest observed effects level (LOEL), or no observed effects level (NOEL) to replace the chronic value in the calculation. It is recognized that estimating the NOEL/LOEL is an inherently uncertain process, however.

Sediment criteria derived from the background or reference area benthic tissue concentrations might serve as useful screening values in the context of the overall assessment. This is based on the presumption that the populations in the reference area are not declining as a function of their exposure to "background" concentrations of contaminants. This also assumes that the reference concentrations are less than the values back-calculated from the chronic toxicity data, the LOEL, or the NOEL.

Risk, as defined above, is fundamentally probabilistic. Therefore, every attempt possible should be made to develop the sediment risk assessment in a probabilistic framework. As the result of such a probabilistic framework, distributions of exposure would be compared statistically with distributions of toxic benchmarks for each species-contaminant comparison. Statistical testing of differences between mean values is an appropriate component of risk characterization. However, statistical measures of overlap of distributions, or estimates of the probabilities of exposure being less than screening values or greater than values associated with unacceptable risk should become standard components of the proposed risk assessment framework.

5. Regional Dioxin Values

A. The dioxin/furan criteria for sediment classification as Category I are based on detection levels, as discussed in Reference 89. The comparative paucity of dioxin and furan toxicity data for species representative of the marine benthos make it difficult to evaluate the efficacy of the 1 and 4.5 ppt criteria. For example, the assumption of using ½ of the detection limits to compute the sum of the non 2,3,7,8,-substituted compounds leads to the 4.5 ppt criterion; clearly, increasing or decreasing this apparently arbitrary value (i.e., ½) would correspondingly increase or decrease the permissible concentration for Category I classification. A dioxin value of 10 ppt (i.e., Category 3 in Reference 89) might prove sufficiently protective, although the necessary supporting toxicity studies should be performed with at least the species of *Nereis* and *Macoma*. Additional studies appear warranted given that the decision criteria were evidently developed on the basis of tissue levels for fish and animals, not including these representatives of the benthos. It is further pointed out (Reference 89) that the many of the pathway coefficients, for example, the trophic level transfer coefficient, were conservatively (i.e., pessimistically) defined in the assessments used to derive the protective criteria for these compounds. More realistic values would, of course, lead to higher permissible concentrations. Cook et al. (1993), cited in Reference 89, suggest a value of 50 ppt as a "low risk" concentration for adverse effects on fish. While additional studies appear needed to justify the classification criteria for dioxins and furans, the information summarized in Reference 89 suggests that a value in the range of 10-50 ppt might be just as logically selected as the current criteria based on detection levels or fractions of detection levels.

B. Ideally, predictions of significant undesirable effects from dioxins and furans would derive from comprehensive, quantitative environmental transport, bioaccumulation, and toxicity assay data. In the absence of these necessary studies, dioxin criteria for sediment classification should at least be developed using quantitative risk assessment methods that emphasize variability and uncertainty in all components of the analyses. Potential differences in exposure among human population subgroups, as well as variability in routes of exposure would logically be addressed in such analyses. The implications of these sources of uncertainty and variability could be effectively explored using Monte Carlo methods, interval analysis, fuzzy arithmetic, or other analytical tools that characterize uncertainty and propagate them through the computations.

Again, the information provided in Reference 89 suggest that benchmarks protective of human health are more conservative than values derived according to ecological risk paradigm. However, it remains difficult to assess the meaning of these comparative benchmarks, given the differences in ecological verus human health endpoints and the assumption that cancer risks and hazard quotients adequately assess "human health."

7. Human Heath Risk, Cancer, and Noncancer

C. The determination of the Rfd's for noncancer health effects was described briefly in the Appendix

for Table 1and summarized in Attachment C. The Rfd values appear to have been consistently derived; the appropriateness of these values is contingent on the usual set of pessimistic assumptions attendant to the standard USEPA human health risk assessment process. More appropriate (i.e., realistic) values might derive from a probabilistic estimation of these Rfd's, wherein distributions or at least ranges of parameter values were included. The Rfd could then be selected, for example, as the 95th percentile of an estimated distribution (e.g., Monte Carlo methods) or an upper bound (e.g., interval analysis).

The whole body/fillet conversion factor of 1.35 appears to have been selected as a mid-point value of the range of 1.2 - 1.5 reported for lipophilic substances in other New York-related studies. To the extent that this value was used for non-lipophilic compounds, bias may have been introduced to estimates of those Rfd's.

An benthic tissue Rfd has been derived for lead as 1.25 ppm (rounded to 1.3) in relation to the East River Project (Reference 88). It was assumed that this approach was applicable for the HARS assessment and the value of 1.3 is listed for lead in Table 1. The Rfd appears conservatively estimated given assumptions (outlined specifically in Reference 89) concerning
patterns of consumption, fish behavior, and trophic transfer efficiency. This analysis also included an additional dietary component of lead; in a sense, the fish consumption pathway was double-counted.

The derivation of the lead Rfd might have proceeded more realistically by defining the parameters as distributions (or at least ranges) and incorporating these uncertainties into the calculations. This refers to not only the fish consumption calculation, but also to the estimates of lead exposure from drinking, water, air, dust, soils, paint, and diet.

D. The hazard quotients for noncancer health impacts might be useful for screening-level calculations in relation to the HARS study. However, these measures are extremely limited in their ability to "predict significant undesirable effects". While quotients less than 1 might suggest some minimal likelihood of health impacts, values greater than 1 provide little information concerning the possible magnitude of impact in the context of exposure(dose)-response relationships. Is a quotient of 2 twice as significant as a quotient of 1? Not necessarily. It depends on the underlying (and usually unknown) dose-response function. The quotient also carries little or no information concerning health impacts that were not specific endpoints (e.g., mortality) in the limited number of toxicity assays that are routinely performed with a small number of species.

8. Ecological Risk

A. The CBR approach has been extensively developed by McCarty and co-workers. Given the common limitations of available data, the methodology seems applicable for estimating tissue levels for PAHs and perhaps other organic contaminants. However, the ~3-fold degree of imprecision (i.e., 40,000 - 120,000 ppb) associated with this analysis might pose a practical limitation in applying the CBR method. Additionally, assessing the ecological significance of certain narcotic endpoints (e.g., immobilization) may prove challenging. If endpoints including decreased fecundity, decreased growth and reproduction, or increased mortality are to be addressed using the CBR approach, it may prove

necessary to include other ecological models (e.g., demographics, bioenergetics) for extrapolating the CBR results to effects on individuals or populations. Additional community or ecosystem-level endpoints might be considered in the overall assessment of sediments in relation to the HARS; however, these endpoints do not seem to lend themselves to analysis using the CBR approach.

If it can be verified that the congeners of chlorinated organic compounds are equipotent with PAHs in terms of narcosis, an additive CBR model might be justified for this class of endpoints. It might prove that the chlorinated compounds are equipotent among chlorinated congeners, but that the chlorinated compounds are generally more toxic than PAHs. In this case, a conversion "constant" might be required before an additive model can be consistently and appropriately applied across both classes of organic contaminants.

B. The BCF approach may be appropriate for determining effects levels for contaminants where the major pathway of exposure is from the dissolved contaminant. For contaminants where food web transfers might prove just important or more important (e.g., PAHs and other organic chemicals with Kow ~ 4 - 6), the BCF approach might underestimate exposure and subsequent bioaccumulation. In addition, Attachment A indicates other varying assumptions (e.g., water hardness, metal speciation, acute:chronic conversions) and approaches in implementing the combined WQC and BCF methodology to derive benthic tissue concentrations.

C. The tissue levels based on BCF's reported for fish should reasonably be adjusted to account for the general differences in lipid content between fishes and invertebrates. Additionally, it should be recognized that lipid content, at least in fish, varies across species, as well as seasonally and in relation to diet and food quality and quantity. Thus, assuming a constant lipid value in deriving benthic tissue levels from fish data can introduce bias and imprecision.

D. Bivalves appear to concentrate metals more rapidly and to a greater extent than Nereis (e.g., Bryan and Langston, 1992). Similar comparisons obtained for dioxin (e.g., Reference 89). Thus, for regulatory purposes based on conservative approaches, using the typically higher BCF values for bivalves would produce lower acceptable metal concentrations in sediments for classification as Category I sediments. It would seem that whenever possible, BCF values should be developed independently for bivalves and polychaetes. Or, the apparent difference in the ability of these classes of organisms to aecium toxic metals (and contaminants in general) should be built in as an uncertainty factor in deriving tissue levels and corresponding permissible sediment concentrations.

Calculations

13. Trophic transfer coefficient

It was stated in the body of the report (p. 12) that the trophic transfer coefficients for metals were

conservatively assigned a value of 1.0. Curiously, arsenic is assigned a value of 3, suggesting the potential for biomagnification, which if justified for any metal would pertain mainly to methyl mercury. However, in the absence of human health or toxicity data, arsenic drops out of the analysis at any rate (i.e., Table 1).

The values of trophic transfer coefficients for PAH's were also conservatively selected (i.e., 0.1, Attachment C). Studies described on p. 14 suggest >90% elimination or metabolism of ingested PAHs. A transfer coefficient of 0.02 was cited between fish and invertebrates, although the value might have been as high as 0.23. The transfer values for pesticides were derived using the Gobas (1993) model, which was developed originally for PCBs. The resulting values ranged from 1 - 2.47 and appear consistent at least with observations of some pesticide biomagnification.

14. Rate of Fish Consumption

The fish consumption rate of 6.5 g/d converts approximately to one meal of 6 oz. of fish every two weeks during one year. The accuracy of applying this number generally across people of different age, size, and geographical location is certainly open to argument. However, it would appear highly probable that certain subpopulations of those who regularly utilize the regional marine resources would characteristically consumer more fish than the 6.5 g/d value. It would certainly be appropriate to include an additional analysis that focused on fishermen and other susbsets of the regional populace that eat more than this default rate. However, this introduces the question concerning whom the sediment classification criteria are meant to protect; identification and characterization of the "stakeholders" in relation to this assessment might assist in refining exposure parameters throughout the entire assessment.

<u>General</u>

15. Assumptions Specific to the HARS

Depending on the resources and time available to conduct an assessment, it would be possible to develop regional or more site-specific values for nearly all the factors that enter into the assessment. Regionally-specific values for all the exposure parameters in the equations listed in Reference 88 and Attachment C could in theory be obtained. (One would hope that the fundamental toxicity of the compounds (e.g., Rfd's) would not vary by region.) Reality obviously imposes constraints on the number of parameters that can be estimated on a regional basis. Therefore, the entire calculus underlying the exposure assessment should become the focus of a comprehensive and detailed sensitivity/uncertainty analysis. The results of such analysis would include the identification and rank-ordering of the input values in terms of their importance in defining sediment criteria for each of the contaminants of concern. Using these results, available resources could be judiciously allocated to obtain regional estimates for the key parameters in the exposure assessment.

Unfortunately, while such analyses of exposure have proven valuable in understanding and refining other risk assessments, sensitivity/uncertainty analyses that have also included the toxic benchmark data have emphasized that the main limitation in risk assessment lies in the paucity of relevant and reliable toxicity data. There is no simple solution to this problem other than acquiring the necessary data. At the same time, these more comprehensive sensitivity/uncertainty analyses can rank the contaminants in order of their probable human health and/or ecological concern. The more critical toxicity data can be identified through this process.

17. Synergistic effects

It is certainly desirable to develop the capability to assess the possible synergistic effects of exposure to multiple contaminants - multiple exposure is the real-world situation. We currently lack the necessary data and toxicological understanding to consistently and reliably predict the impacts of exposure to multiple contaminants. The additive model appears to work for certain classes of compounds, as suggested by the equipotency observations in McCarty's work. At the same time, there are repeated instances of the failure of the additive model. Certainly, if one of the contaminants is more acutely toxic than others, it will likely "mask" the effect of the less malevolent compounds and additivity will not be observed. In other instances, the presence of one contaminant can increase the effectiveness of other co-contaminants. Unfortunately, we by and large lack the models to quantitatively predict from among these possible alternatives.

While research continues to address synergistic effects, it seems prudent to at least continue with general application of the linear model in developing an overall site-wide assessment of risk.