Guidelines for Evaluating Existing Analytical Data to Determine Suitability for Use in Ecological Risk Assessments

Author: Ed Corl
Organization: Atlantic Division, Naval Facilities Engineering Command

Abstract

It is crucial that existing data are evaluated for suitability prior to being incorporated into the ecological risk assessment (ERA) process. Much of the guidance for determining data useability for risk assessment purposes has been developed by the United States Environmental Protection Agency (USEPA, 1992). Data used in ERAs can represent a variety of environmental media (e.g., surface water, groundwater, sediment, soil, and tissue), obtained using a variety of collection and analytical procedures. Basic evaluation of data quality and useability should include consideration of field sampling methods, analytical procedures, detection limits and data quality (e.g., quantitation limits, qualifiers, codes, and blanks). Further considerations of data analysis and summarization procedures should also be made prior to using data in ERAs. For existing data, one should consider the age of the data for appropriateness and the data source. Data should also be evaluated for relation to site activities, sample characteristics (e.g., sample depth, filtering of metals, composite vs. discrete samples, spatial coverage), data quantity and report summaries (i.e., statistics). As with any ERA, uncertainties must be addressed and should include those associated with data evaluations and considerations stated above.

Issue Discussion

Introduction

This issue paper discusses procedures for evaluating existing analytical data to determine if they are suitable for use in an ecological risk assessment (ERA). Available analytical data of interest include those from samples of abiotic media (surface water, groundwater, sediment, and soil) and biotic media (tissue residues). If data collection and analysis procedures were not developed with quantitative risk assessments in mind, the resulting analytical data should be carefully evaluated to determine how (or if) they can be used in a quantitative ERA. Regardless of whether or not existing analytical data can be used quantitatively in an ERA, such data might still be useful for such purposes as guiding additional data collection activities (in terms of analytical parameters, locations, number of samples, etc.) or to qualitatively look at trends in concentrations over time. It should be noted that the procedures contained in this issue paper are intended only as a general guide since each individual site will be different. Professional judgement is important to determine how (or if) to apply these procedures at a particular site.

Evaluation of Data Quality and Useability

The U.S. Environmental Protection Agency (USEPA) has developed guidance for evaluating the useability of data for risk assessment purposes (*Guidance for Data Useability in Risk Assessment (Part A)*, Office of Emergency and Remedial Response, Publication 9285.7-09A, 1992). Although developed primarily to support human health risk assessments, this guidance is also useful for evaluating data useability issues for ERAs. The reader is referred to The EPA guidance document for more details on data useability issues.

Consideration should be given to the following points when evaluating data quality and useability issues for an ERA:

1. **Sampling Methods.** The sampling methods used to collect the data should be appropriate and the sampling design should be sufficient to meet the basic data objectives. Caution should be used when pooling data from different sampling events and/or data based on different sampling methods. Different physical procedures and/or temporal differences between the sampling events may produce varying analytical results. For example, concentrations in surface water from samples collected during two discrete events in different seasons may reflect seasonal influences; pooling these data may mask these seasonal factors, especially if sample sizes between events are very disparate (pooling the data would thus "weight" one event over the other). Different sampling methods (e.g., a Ponar versus an Ekman dredge for sampling sediments) may sample slightly different depths (e.g., 0 to 6 inches versus 0 to 4 inches) which may influence the chemical concentrations found. While it may still be appropriate to pool these data, this should be done only after considering the possible implications to the evaluation.

2. **Analytical Methods.** All analytical data to be used in an ERA should have been generated from currently accepted methods, or from methods of comparable quality. Failure to do this could result in Data Quality Objectives (DQOs) not being met and possible rejection of the data set by the regulators.

3. **Detection Limits.** The detection limits from multiple sets of data should generally be comparable if data are to be pooled. If detection limits are higher than screening values (e.g., Ambient Water Quality Criteria), these data can be considered in the ERA although chemicals with detection limits exceeding screening values are generally carried forward into the BERA (even if non-detect). Data with elevated detection limits are generally most useful in screening ERAs (where maximum concentrations are used) but their use in baseline ERAs (where mean concentrations are used) may be problematical since mean concentrations (which are usually based on one-half of the detection limit for non-detect samples) may be artificially elevated.

4. **Data Quality.** The quality of the data should be evaluated with respect to parameters such as sample quantitation limits, qualifiers, codes, and blanks.

- All data used quantitatively in the ERA should be validated by a qualified data validator using acceptable data validation procedures. When validated

analytical data are limited or lacking, unvalidated data may be considered in the ERA on a qualitative basis. Previously unvalidated data can be post-evaluated by a qualified validator to determine the extent of the data quality.

- Data rejected during data validation (R flag) should not be used in the ERA. Depending upon the nature of the rejection, the data may still be used qualitatively. For example, data can be rejected due to sample holding times being exceeded. However, if the sample results showed a high concentration of a particular contaminant then the severity of the holding time exceedence becomes secondary to the presence of the contaminant.

- Data qualified with a B flag for blank contamination should generally be treated as non-detect (i.e., the same as U-flagged data). Region III is the only EPA Region with a data validation qualifier (B) for blank contamination. It is applied when the sample concentration is less than five times the blank concentration or less than 10 times the blank concentration for common laboratory contaminants (e.g., acetone). Data validators in all other EPA Regions also use the 5x/10x rule, but apply the U qualifier (not detected) rather than the B qualifier. Since there is not a standardized data flag qualifier list, laboratories may use a specialized set of data qualifiers. Therefore, end users of data must have a clear understanding of the specific lab qualifier identification as it will affect the practical utility of the flagged data.

- Data with J (estimated), K (biased high), or L (biased low) flags are acceptable for use in an ERA to represent detected concentrations. K and L flags are unique to Region III and while they represent valid concentrations from a data validation standpoint, conclusions based on such values should carefully consider the associated bias (e.g., a conclusion of acceptable risk may not be appropriate based solely on L-flagged data). Potential issues associated with L and K flagged data are rarely significant since these flags are not generally widely used (data which are too biased will fail the validation criteria and be rejected). J flagged data also indicate that a low or high quantitative bias exists or that the compound was a tentatively identified compound (TIC) if GC/MS was used for the analysis.

Data Analysis and Summarization Procedures

The data that are deemed of sufficient quality for use in an ERA should be screened again to determine which data are applicable to a particular analysis. Not all data collected at a site, even when of sufficient quality, will be useful to an ERA. Consideration should be given to the following points:

1. **Age of the Data.** Since ERAs typically evaluate (current or existing) conditions, only data from the last round of sampling conducted at a site for each chemical group and location should generally be considered unless temporal trends are of particular interest (e.g., for an evaluation of natural attenuation). For groundwater, surface water, and tissue residues, samples from the most recent one-year period should be

considered when evaluating baseline (current) risks to account for potential seasonal variation. For surface soil and sediment, data older than one year are often useable for chemicals that do not readily degrade (e.g., PCBs and metals), although only the most recent samples should probably be used to evaluate baseline (current) risks if there are multiple rounds at the same location since these data best reflect current concentrations. Older data can be used in the screening ERA if they are the only available data but consideration should be given to collecting additional samples if the ERA process progresses to the baseline ERA (Steps 3 - 7).

2. **Data Source.** Data from temporary groundwater wells and field test kits should not be used quantitatively in an ERA since these data/methods do not generally meet the data quality objectives for a quantitative ERA.

3. **Relation to Site Activities.** Surface soil or sediment data collected prior to any major physical disturbance (such as capping or dredging) should not be used in the ERA (other than retrospective ERAs that evaluate past conditions for purposes such as natural resource injury determinations) since such disturbances significantly change the conditions at the site and the samples would no longer represent current conditions.

4. **Sample Depth.** For surface soil, samples collected from depths of 0 to 6 inches (0 to 15 cm) should be used preferentially since they represent the most likely exposures to most ecological receptors. Data from slightly deeper depths (0 to 12 inches; 0 to 30 cm) are also generally useable, especially if surface soil data from 0 to 6 inches are limited. Caution should be used when mixing data from different depth strata since site-related chemical concentrations frequently vary with depth. In some situations (e.g., burrowing animals are a key receptor), deeper soils (at depths of up to five feet below ground surface) may need to be considered on a site-specific basis, depending upon the assessment endpoints selected and the nature of the conceptual site model.

   For sediment, samples from depths of 0 to 6 inches (0 to 15 cm) are generally preferred based on likely ecological exposure potential. Samples from depths of 0 to 12 inches (0 to 30 cm) are also generally useable, especially if data from shallower sediment strata are unavailable or limited, or if the redox boundary (which typically represents the bottom of the biologically active zone) is deeper than 6 inches.

5. **Filtered Metals.** For surface water and groundwater samples, total (unfiltered) metal concentrations are generally used during the initial screening assessment steps although filtered data are useful in the baseline ERA. It is generally accepted that filtered metal concentrations more closely represent the biologically available portion of metals in water.

6. **Composite Samples.** The use of composite samples, although sometimes necessary (e.g., to obtain sufficient sample weight for tissue analyses), should be used with caution especially if the data will be analyzed statistically. In general, data obtained from composite samples should not be combined with data from grab samples because important information about the sample variability is lost in composite samples. If samples over relatively large depths or intervals are required (e.g., water

samples with depth in a relatively deep water body), discrete samples over limited depths or intervals are preferred to composite samples.

7. **Quantity.** Data should be evaluated (for each medium of interest) to determine if the quantity of available data are sufficient to perform the ERA. In some cases, sufficient data may be available to start the ERA process (screening steps) but more data would be needed to complete later steps of the ERA process.

8. **Spatial Coverage.** Available analytical data should be evaluated to determine if they are representative of the site and if there is sufficient spatial coverage in all habitats and areas of potential interest to the ERA. While this is critical for performing a baseline ERA, a screening ERA may be completed using data from a more restricted spatial area as long as all areas and general habitat types (e.g., wetlands) where a release is likely to have the greatest effects have adequate coverage. Consideration should be given to how data will be grouped (e.g., upgradient/downgradient, by habitat or individual area) in the ERA when making the determination of adequate data quantity and spatial coverage.

9. **Data Summaries.** When summarizing data, the following should be considered:

    - The maximum detected concentration (or maximum detection limit if both sample duplicates are non-detects) should be used from sample duplicates in the screening ERA. In the baseline ERA, other measures such as averaging sample duplicates or calculating the 95 percent upper confidence limit (UCL) of the mean of the two duplicates can be considered.

    - Non-detects should generally be treated at one-half of the sample detection limit when calculating data summary statistics during the baseline ERA. However, individual samples should be excluded from calculations if the sample detection limit is more than five times higher than the method detection limit and the sample result is labeled non-detect (U or UJ flag). Excluding these data points should only be done if this procedure is included in the work plan and agreed to by all parties.

    - When calculating exposures in the baseline ERA, the mean calculated concentration (or the 95% UCL of the mean if that statistic is used) should not be used if it exceeds the maximum detected concentration; the maximum should be used in this case.

10. **Uncertainties.** Data gaps, data quality issues, and analysis methods should be discussed, as appropriate, in the uncertainty section of the ERA.

Summary

Many factors influence the suitability and adequacy of existing analytical chemistry data for use in an ERA. In some cases, the step of the ERA (screening versus baseline) will dictate the suitability and/or adequacy of existing data. The way data are summarized and used in the ERA can also be influenced by the step of the ERA process.

General guidance for evaluating the useability of data for risk assessment purposes can be found in the USEPA document *Guidance for Data Useability in Risk Assessment (Part A)* (Office of Emergency and Remedial Response, Publication 9285.7-09A, 1992). The NFESC IR QA Program is another potential source of guidance related to this issue (contact Pati Moreno at 805-982-1659, e-mail: morenop@nfesc.navy.mil).

Point of Contact

Ed Corl
Atlantic Division
Naval Facilities Engineering Command
Code 18328, Technical Support
1510 Gilbert Street
Norfolk, VA 23511-2699

Acronyms

ERA – Ecological Risk Assessment
DQO – Data Quality Objective
IR – Installation Restoration
NFESC – Naval Facilities Engineering Service Center
PCBs – Polychlorinated Biphenyls
UCL – Upper Confidence Limit
USEPA – United States Environmental Protection Agency

Glossary

Abiotic media: an environmental medium of non-living substances (e.g., sediment, water and or soil) that can function as a compartment for environmental contaminants
Biotic media: an environmental medium comprised of living materials (e.g., animal tissue, blood and/or whole body) that can be a compartment for environmental contaminants.
Composite sample: a sample of environmental media (e.g., sediment, surface water, groundwater) which is comprised of several individual, discrete samples taken from a defined spatial arrangement and through similar procedures.

Data quality objectives (DQOs): qualitative and quantitative statements that define the type, quality, and quantity of data necessary to support defensible risk management decision-making. used to develop an effective sampling plan which avoids the collection of data that are inconsequential

Data validation: the process of verifying and qualifying environmental data so that they are consistent with, and exhibit characteristics of the standards required for an intended use (e.g., ecological risk assessment).

Ecological risk assessment (ERA): process that identifies stressors (e.g., chemical, physical) that may alter ecosystems and quantifies the probable severity of adverse effects on those ecosystems

Filtered metals: a surface water sample which is processed through a filter in order to eliminate any particulate to which metals may be adsorbed; only dissolved metals remain in the environmental sample.

Human health risk assessment (HHRA): process that identifies stressors (e.g., chemical, physical) that may affect human health and quantifies the probable severity of adverse effects on humans.

References

U.S. Environmental Protection Agency (USEPA). 1992. *Guidance for Data Useability in Risk Assessment (Part A.*, Office of Emergency and Remedial Response, Publication 9285.7-09A, 1992.