ANL/CP--71902

DE92 003418

Addressing Data Heterogeneity:
Lessons Learned from a Multimedia Risk Assessment

Halûk Özkaynak, Jianping Xue, and David A. Butler
Harvard University
79 John F. Kennedy Street
Cambridge, MA 02138

Lynne A. Haroun, Margaret M. MacDonell, and Donald J. Fingleton
Argonne National Laboratory
9700 S. Cass Avenue
Argonne, IL 60439

For Presentation at the 84th Annual Meeting of
the Air and Waste Management Association
Vancouver, British Columbia

June 16-21, 1991

INTRODUCTION

Cleanup activities are currently being conducted by the U.S. Department of Energy (DOE) at a former chemical plant site that has been inactive for more than 20 years. The Army produced nitroaromatic explosives at the 220-acre site during the 1940s, and radioactive materials of the uranium and thorium series were processed there by DOE's predecessor agency during the 1950s and 1960s. The chemical plant site includes about 40 buildings (some of which contain process vessels, pipes, and sumps), four waste pits that contain sludge and surface water, two contaminated ponds, two solid waste dump areas, a former coal storage area that served the power plant, and various localized spill areas. In 1989, the U.S. Environmental Protection Agency (EPA) listed the site on its National Priorities List. Hence, site cleanup is being conducted consistent with EPA's Superfund process.

Chemical and radioactive contaminants are present in soil, surface water, sediment, and groundwater at the site as a result of both past releases and disposal activities and subsequent contaminant migration. Samples have been collected from these media over a number of years under both DOE's environmental monitoring program and the site characterization program of the Superfund process. Results of sample analyses have been compiled in a computerized data base. The current data base for the site contains more than 120,000 records, each of which includes a number of different entries (e.g., contaminant, medium, location coordinates, depth, and analytical method).

These data are being evaluated in the context of potential exposure pathways that are currently present at the site or that may be present in the future, in order to estimate possible adverse impacts to human health and the environment in the absence of cleanup. This evaluation of short-term and long-term "baseline risks" is an important component of EPA's Superfund process because it serves to focus the selection of cleanup remedies. Estimates of site-related risks are presented in a baseline risk assessment report, the multiple objectives of which are to:

- Estimate the possible outcome of not remediating the site -- considering both discrete source areas and general site-wide contamination -- by quantifying health impacts associated with possible exposure pathways such as inhalation, incidental ingestion, and dermal absorption under both possible current scenarios and hypothetical future scenarios. Environmental impacts are also assessed, but this paper focuses on the human health component of the baseline risk assessment.

- Identify data gaps where they exist so that additional information can be collected to support the cleanup decision for the site -- e.g., by sampling anew for contaminants that may be present based on historical records and for which toxicological data suggest a possible health concern, but for which no characterization data are available.

- Define specific areas of the site for which cleanup should be considered based on estimates of current and future risks -- e.g., the extent of a former waste dump or spill area for which capping or excavation may be appropriate to reduce potential exposures.

- Identify areas that do not require remediation -- e.g., those not likely to result in incremental adverse health impacts greater than targets identified by the EPA as "acceptably protective" for Superfund sites. These targets are: an incremental lifetime carcinogenic risk of 10^{-4} to 10^{-6} and a hazard index of less than one for noncarcinogenic health effects (determined by summing the ratios of estimated daily intakes by standard EPA reference doses for the appropriate contaminants of concern).[3,4]

- Establish preliminary cleanup criteria -- e.g., specific concentration levels for the primary risk-driving contaminants, to minimize residual risks at the site following cleanup.

- Provide a baseline for comparing the protectiveness of candidate cleanup alternatives -- e.g., to support a decision between capping or excavating a dump area or leaving it as is.

- Focus the remedy selection process for the site -- e.g., by highlighting areas and contaminants that pose the primary health and environmental concerns.

The preparation of a baseline risk assessment for the site to meet these objectives required an extensive review and interpretation of the "raw" data to ensure that the estimated risks and subsequent cleanup decisions would be based on appropriate values. Most of these investigations involved characterization, validation, and analysis of diverse and heterogeneous pollution measurements gathered from different media. Collectively, these assessments attempted to address many of the technical "data heterogeneity" problems associated with the site-specific environmental data. Various specific tasks were addressed as part of this effort, certain of which are interrelated; these tasks included:

1. Verifying and validating site data, working with both the hard copy and the electronic data base and with supporting laboratory sheets, to minimize the use of erroneous values in the risk assessment and to identify additional sampling needs (e.g., check samples).

2. Compiling data from both random and biased sampling programs to determine whether the site was adequately characterized to support the baseline risk assessment and subsequent cleanup decisions (taking into account historical information) and to interpret certain values that may be skewed high because of an intentional sampling bias.

3. Applying summary statistics and sorting programs to gain a broad understanding of the site problem and to identify outliers, for which subsequent data checks and discrete hot-spot analyses could be appropriate.

4. Reviewing both the results of sampling at local areas off-site and the available literature to place measured concentrations in perspective with background values for certain naturally occurring contaminants.

5. Addressing non-detects in the data base -- considering the specific contaminant, medium, and detection limit; its detection frequency and natural occurrence (or anthropogenicity) in that medium at the site; and historical information -- to assess the likelihood of its presence and to identify an estimated concentration value for use in the risk assessment, as appropriate.

6. Developing strategies for outliers and hot-spot analyses, considering verification/validation and check samples, reviewing historical information, and targeting discrete locations as representative of those that could result in reasonable maximum exposures for related pathways at the site.

7. Evaluating the results of surface and subsurface sampling, considering historical information and the potential impact on future exposure estimates for various media.

8. Conducting location-specific analyses for soil and groundwater -- i.e., borehole-by-borehole and well-by-well analyses, respectively -- to address the heterogeneity of contamination across the site and to bound the range of estimated impacts.

This paper discusses the methodology used to address each of these tasks and the lessons learned during the assessment process. Statistical issues and recommended future directions for dealing with technical aspects of this project and with similar multimedia risk assessment projects are addressed in the final discussion.

METHODOLOGY

The approaches taken to address each of the data-management tasks identified for preparing the risk assessment are described individually below. Results of each effort are discussed in the following section.

Data Verification and Validation

Consistent with the quality assurance/quality control considerations for data at a Superfund site, the data base was intensively reviewed to assess the adequacy of existing values and of the available data base as a whole to ensure that subsequent calculations for the risk assessment were based on defensible information. The intent of the verification effort was to catch mistakes that were made on-site while compiling the analytical results received from various laboratories into a single data base. The intent of the validation effort was to identify mistakes made in the laboratory, e.g., while entering information onto the original data sheets.

To perform this task, data were sorted a number of different ways to identify problems such as double-counting because of duplicates, split-counting because of misspelling an analyte, and inconsistent treatment of non-detects (e.g., using "zero", a hyphen, the contract-required detection limit, or the method detection limit without explanation). The various sorts included the following combinations: (1) analyte, identification number (laboratory/site), and concentration -- to identify misspelled analyte entries; (2) analyte, identification number, and location -- to identify mistakes in entering location numbers, typically substituting "O" for "zero" (which, when corrected, facilitated subsequent sorts and statistics); (3) analyte, identification number, location, depth, and concentration -- to identify duplicates; (4) analyte, medium, detection limit, and concentration (especially zero) -- to identify inconsistencies in reporting non-detects; (5) analyte, identification number, medium, and detection limit -- to identify mistaken linkings of certain contaminant concentrations with certain media (e.g., reporting values for soil that should have been for surface water); and (6) analyte, moisture content, medium (e.g., soil or sludge), and concentration -- to identify samples for which corrections had or had not been made for consistent reporting on a dry-weight basis.

The data were also sorted by analyte, identification number, and concentration to identify the highest values for each contaminant in a specific medium (primarily soil), which could often be described by a "top twenty" hit list. These values were then compared to others measured for that contaminant in that medium, considering sampling location and site history, results for other depths and/or for nearby samples, and typical background concentrations (as appropriate). As for the other sorts, those values identified as seemingly inconsistent with other information were targeted for verification and, as appropriate, for validation -- e.g., where no verification error was found.

Random and Biased Sampling

The objectives of the various data collection efforts at the site were considered in interpreting the results of those efforts and determining whether they could be combined for the risk analyses. For this task, sample points were plotted on a site map using different symbols for each of the three major sampling studies. Field measurements and historical records were then reviewed to determine whether any critical data gaps existed, i.e., whether uncertainty in the risk calculations could be significantly reduced by collecting additional data. In cases where measurements appeared inconsistent, for example where adjacent samples were considerably different, archived samples were reanalyzed (e.g., for radionuclides) or additional samples were taken (e.g., for chemicals).

Summary Statistics

The data were statistically analyzed and summary statistics were generated to present a general overview of characterization results and to assess the type(s) of distribution to support subsequent statistical manipulations (e.g., to determine the appropriateness of log-transforming the data for manipulation). For this task, the data were sorted by contaminant and medium so that detection frequencies could be determined. Location identifiers were included in the output so that the results could be reviewed in concert with other information such as historical records to identify potential data gaps and to assess whether certain contaminants typically coexisted at the site.

Background Comparisons

For this task, concentration data were plotted with depth for the off-site area at which background soil samples were collected. Because background data for nearby surface water and groundwater were limited, available information was supplemented with information from the literature; the literature was also reviewed for site-specific soil parameters such that the body of site characterization data could be put in perspective with the range of values commonly found in both the state and the country for these parameters in various media.

Non-Detects

Having identified the various presentations of non-detects in the data base as part of the data verification/validation effort, these records were standardized and then grouped according to contaminant and medium in order to apply different strategies for substituting values for non-detects to determine exposure point concentrations for the risk assessment. The substitutions for non-detects were contaminant-dependent and were based on the results of statistical summaries and background comparisons, as well as historical information. For example, various runs were made to assess the impact on risk calculations of replacing non-detects with the full detection limit, half the detection limit, and zero. Based on these and other more comprehensive investigations, a simplified strategy was selected to fill in values for the non-detects.

Outliers

The outliers -- e.g., the unusually high values reported for certain contaminants -- that were identified on the basis of sorting and statistical analyses were plotted on a site map to focus data verification/validation and data gap analyses and to indicate locations at which hot-spot analyses could be appropriate. In addition, the site soil was visually inspected in some cases to determine whether the value could be a "reasonable" anomaly (e.g., a very high lead measurement was due to a hunter's buckshot); for other anomalies, additional samples were taken.

Surface and Subsurface Sampling

Based on historical information, subsurface contamination was expected at discrete locations on-site. Following closure of the Army ordnance works, process lines and contaminated soil were excavated as part of an extensive decontamination and dismantlement effort. In addition, the site was regraded prior to construction of the processing facility for radioactive material, with cut and fill operations sometimes covering existing surfaces with 8 to 10 feet of soil. Therefore, most of the contamination at the site is surficial and derives from the uranium processing facility, but limited subsurface chemical contamination is present as residual material that was missed during the previous cleanup effort. To assess the nature of subsurface contamination at the site, soil was sampled at various depth intervals as part of both the biased and random sampling programs. For the biased program, blueprints of the former ordnance works and old cut-and-fill records were reviewed to focus the subsurface sampling for chemical contaminants. The data were then sorted with a depth record to identify locations at which subsurface concentrations were elevated above available soil cleanup criteria, risk-based concentrations (using a screening level concentration-toxicity approach), and/or background levels. Concentration profiles of various contaminants were also compared to identify similarities or differences in distribution with depth across the site. Subsurface locations at which certain contaminants are elevated were also mapped to permit an overlay with site-specific geological and hydrological data to support the estimate of potential leaching to groundwater.

Location-Specific Analyses

Exposure and risk estimates were generated for each sample location for surface soil and for groundwater. This approach was selected to address the heterogeneity of contamination at non-source areas on-site in order to provide a range of risks associated with different areas and to highlight locations to be considered in the cleanup decision. For surface soil, borehole-by-borehole analyses were conducted and the results were plotted by location on the site map.

In contrast to typical situations, groundwater contamination at the site is not uniform because the chemical plant rests atop a groundwater divide, source areas are unevenly distributed, and certain flow paths are preferential. In addition, the various monitoring and characterization wells were completed at different depths (e.g., in both the shallow and deep aquifer), rendering the calculation of a 95% upper confidence limit of the arithmetic average concentration for contaminants in groundwater across the site inappropriate. Therefore, groundwater was analyzed on a well-by-well basis to estimate risks associated with drinking water ingestion under a hypothetical future resident scenario, and the results were plotted on the site map.

RESULTS AND DISCUSSION

The baseline risk assessment for a hazardous waste site can play a significant role in the decision-making process for site cleanup, as described in the multiple-objective list presented in the Introduction. Therefore, the defensibility of the information presented in this assessment is important to the cleanup decision. In addition, a sound data review and interpretation effort can strengthen community relations that are often strained because of a perceived risk problem and a common feeling that "the more the better, regardless of cost" should be the cleanup anthem. A clear presentation of data and risks in the baseline assessment can make a strong argument for subsequent decisions that weigh the benefits of incremental risk reduction against the engineering costs. Therefore, it is essential that the risk assessment be developed in a technically defensible manner, beginning with data management.

Two main factors are important in defining and carrying out the data-management tasks associated with preparing a baseline risk assessment for a hazardous waste site. These interrelated factors are site history and the adequacy of available data. Superfund sites vary considerably in the nature and extent of contamination present and in the availability of characterization data. For the site addressed in this paper, eight tasks were identified as components of the data interpretation effort to support risk calculations. It is expected that these tasks could be applied to most hazardous waste sites. General lessons learned from implementing the individual tasks are presented below.

Data Verification and Validation

Data verification and validation is the most time-consuming of all the tasks, but it is crucial to the selection of appropriate values for the risk assessment. The most important lesson learned from this task actually addresses the setup of the original data base rather than its review. Another lesson is that data should be interpreted carefully at the time it is received, to preclude the need to track down laboratory data sheets at some future date when it is much more difficult to evaluate errors.

To efficiently interpret and manipulate a data base (e.g., run statistical procedures and sorting exercises), computerization is essential. It would have been almost impossible to manage the extensive amount of data generated for the site in question unless it had been in an electronic form. However, data must be carefully transcribed from the laboratory sheets into the computer and should be verified at the time of transcription (e.g., with double-key entry) or verification of a very large data base during the assessment phase can take months. The correction of verification mistakes is straightforward and does not typically result in the rejection of data as unusable for the subsequent risk calculations. Examples of "verification" mistakes that were identified as part of the data review are the misplacement of decimal points and the misreading of units such that values reported in the data base were orders of magnitude higher than those actually measured and recorded in the laboratory.

Compared with data verification, data validation is much more difficult and must often rely on detective work and expert laboratory judgment. Examples of "validation" mistakes that were made include (1) the misidentification of an analyte, e.g., confusing the notation for arsenic (As) with that for aluminum (Al) such that the value reported for "arsenic" in the data base was extremely high; (2) the misidentification of the medium sampled, e.g., writing "water" instead of "sludge" -- which also led to orders-of-magnitude mistakes for certain contaminants; and (3) the failure to note that a sample analyzed for nitrate had in fact been preserved in nitric acid. Validation corrections are usually much more difficult to make than verification corrections, and the data can often be deemed unusable. Another lesson learned from this task is that the expert judgment that is invoked during this process is needed to provide a reasonable explanation if high values are rejected, or this rejection might otherwise be construed as the arbitrary and capricious exclusion of outliers that could be considered to represent hot spots.

Random and Biased Sampling

By reviewing the random and biased sampling results and plotting all locations sampled for specific parameters or classes thereof (e.g., metals or organic compounds) we were able to more easily assess data gaps. On the basis of this review, it was determined that the site characterization programs had adequately addressed areas at which residual ordnance-works contamination might be present. It was also determined that the results of the various sampling programs could be combined for the purpose of the baseline risk assessment because the biased sampling at known source areas would tend to result in the use of conservative values for exposure point concentrations for the site-wide analysis.

Summary Statistics

The lesson learned from applying statistical procedures to the data is that the verification and validation effort should be completed prior to designing additional sampling efforts or running the calculations for the risk assessment on the basis of summary statistics. That is, because these efforts were being conducted in parallel at the site due to schedule constraints, summary statistics had to be rerun and reinterpreted (e.g., relative to outlier contribution to hot spots) several times due to changes in the data base as the verification/validation results were incorporated. Statistical information can be used for a number of purposes: getting a general picture of the multimedia site problem, identifying outliers, and providing a first step toward streamlining the assessment by focusing the list of contaminants considered. These applications are discussed below.

The results of statistical manipulations were used to present a general overview of site contamination in the risk assessment. The following information was tabulated for each analyte in each medium as part of this presentation: (1) the range of values measured (the lower limit of detection to the maximum detected); (2) the detection limit, detection frequency (as a ratio of the number of detections to the number of valid samples), and the percent of detects; and (3) the concentration of that analyte in blank samples (e.g., field and laboratory blanks). This information was used to identify data gaps by reviewing the compiled data and comparing the measurements with historical records so that additional samples could be collected as needed in time to support the risk assessment and site cleanup decision. For example, where the number of analyses for a certain organic compound in surface water was small and its actual presence was in doubt (based on both historical information and measured concentrations in the samples and in attendant blanks), additional samples were targeted.

The data were also used to plot frequency histograms, which were used to determine that in fact contaminant concentrations in site soil were log-normally distributed. Statistical information was also used to determine the upper 95% confidence limit of the arithmetic average (C_{95}) for contaminants in site soil to be used as representative concentrations for estimating general site-wide exposures. In addition, outliers were identified from the statistical analyses, and the high values were commonly subjected to verification/validation to ensure that the value was usable in the risk assessment. Those values deemed acceptable for use in subsequent risk calculations were addressed via location-specific analyses to estimate hot-spot exposures.

The results of statistical analyses were also used to support the selection of contaminants of concern in the baseline risk assessment. That is, because a considerable number of analytes had been detected in various media at the site, it was considered important to focus the risk assessment by excluding those that contributed insignificantly to preliminary risk estimates, as supported by other analytical, toxicological, and/or historical information. For example, analytes considered common laboratory contaminants that were present at less than 10 times the maximum concentration detected in blanks were considered for screening from further analyses (e.g., acetone). Similarly, analytes for which the detection percentage was very small (e.g., less than 3%) were also considered candidates for screening. Prior to excluding such analytes from quantitative analyses in the risk assessment, calculations were made on the basis of a concentration-toxicity screen to assess their relative contribution to site risks.

Background Comparisons

Highlighting a comparison of measured values with background concentrations for naturally occurring substances can strengthen the general presentation of the site problem and the ultimate risk management decision. For example, because the dermal absorption of metals appeared to drive estimated risks for the site in question, it was helpful to explain that the results were considerably affected by uncertainties in the toxicological data and that in fact the on-site levels were within the range of local background values, for which estimated risks were comparably high. Similarly, if ingestion of lead from soil or arsenic from groundwater appeared to drive risks, it would be helpful to place the levels in perspective with common background values. In addition to alleviating a perceived risk problem and clarifying the issues (such as uncertainties in route-to-route extrapolations for the toxicological data), the results of this comparison can be used to identify data verification/validation needs. As an example, only one sample of mercury out of 25 taken at the off-site background location exceeded the detection limit in soil, and the reported value was extremely high; this flagged the need for further data review. The analysis of background values can also be used to support surface/subsurface comparisons, e.g., for metals. Finally, the results of statistical analyses of background data can be used during the subsequent treatment of non-detects, e.g., to identify replacement values that could be appropriate.

Non-Detects

By applying various approaches to substitute values for non-detects in the data base, we learned that a single approach is inappropriate for the variety of multimedia contaminants at the site. Historical information and detection frequencies were used to select replacement strategies for the various non-detects. For example, certain organic compounds are not naturally present in soil and were generally absent from the site because of limited use during past operations (as indicated by a low detection frequency and our spatial data plots). For organic compounds, non-detects were replaced with a zero if the compound was never measured at the site, and replaced with half the detection limit if a measurement above the detection limit existed. Little difference resulted from replacing non-detects with either the full detection limit or half the detection limit for radionuclides, metals, and inorganic compounds that are naturally present in soil -- primarily because the detection frequency was usually high. Hence, for consistency, non-detects for metals and inorganic ions were replaced with one-half the detection limit. However, because radionuclides are the primary contributor to risks via the various exposure pathways, the full detection limit was used for these parameters for the sake of conservatism. Non-detects in groundwater were addressed on a well-by-well basis so that site-specific hydrogeological conditions could be incorporated (e.g., the presence of the groundwater divide and of discrete source areas). Additional discussion of this issue is presented in the final section of this paper.

Outliers

The lesson learned from the outlier review task is that outliers should be identified and plotted early in the data management process so results can be used to focus additional sampling and identify hot-spot areas for subsequent analysis. In presenting the overview of site contamination in the risk assessment, certain outliers were excluded from the summary statistics and identified separately (e.g., in footnotes) such that the general picture of contamination would not be skewed. However, these values were emphasized in location-specific analyses to identify areas of the site for which cleanup should be considered on a risk basis. For certain parameters, the outlier results and the subsequent assessment of data gaps led to an expanded sampling program (e.g., for groundwater), relative to both analytes and sampling frequency. By plotting the high values for each contaminant, locations at which cleanup could be considered were highlighted; in most cases, these areas were consistent with expectations based on historical information. For example, the high values for polychlorinated biphenyls were plotted at locations that formerly housed electrical equipment and those for polynuclear aromatic hydrocarbons indicated that they coexisted at one location -- the former coal storage area. However, other contaminant plots identified areas adjacent to buildings and to the waste pits at which spills had occurred that would not necessarily have been targeted for cleanup on the basis of available historical information. These plots were used to target additional areas at which the use of engineering controls and worker-protection equipment would be considered during future activities (e.g., excavation) to minimize potential exposures.

Surface and Subsurface Sampling

We found that a clear presentation (e.g., with plots) of surface/subsurface data can be used to streamline the risk analysis. That is, if the general extent of the problem (limited in our case) is appreciated early in the risk assessment process, the consideration of subsurface areas can be tailored to meet focused objectives, e.g., by concentrating on a few pockets with elevated contaminant concentrations rather than on the entire subsurface. The results of this effort were used to support the decision on what, if any, subsurface contamination should be removed as part of site cleanup.

For radioactive contaminants, results were used to estimate external gamma exposure levels and related risks at the site and to identify subsurface locations at which concentrations exceeded regulatory levels. Although cleanup criteria are not generally available for residual chemicals in soil, standards have been identified for residual radium and thorium. To consider a DOE site for use without radiological restrictions, applicable DOE standards and EPA requirements that are considered potentially relevant and appropriate identify the following criteria: 5 picocuries per gram (pCi/g) above background averaged over 100 m^2 for the surficial 6 inches of soil and 15 pCi/g for each 6-inch subsurface increment.[3,4] Hence, the surface/subsurface presentation supported the regulatory compliance effort for the site.

Subsurface locations at which certain contaminants are elevated were also mapped to permit an estimate of potential leaching to groundwater that reflected actual site conditions. Although literature values are often used to estimate leaching to groundwater, it is much preferred to use site-specific information where available. In fact, leaching rates at the site are influenced by location because geotechnical and hydrological conditions vary across the 220 acres. For example, the presence/absence and thickness of different stratigraphic units and the estimated magnitude and direction of flow velocities are location-dependent. Hence, an appropriate understanding of the spatial orientation of subsurface contamination can reduce uncertainties in the parameters used to estimate the number of years it may take for the contaminants to leach to groundwater and the resultant concentrations that may be subsequently ingested. This strategy improves the estimate of potential future risks from drinking-water ingestion for the site.

In addition, areas mapped with elevated subsurface contamination were used to focus the "excavation and redistribution" scenario for a hypothetical future resident. That is, this subsurface soil was assumed to be brought to the surface via excavation for the basement of a hypothetical house under the future scenario for the site. The "basement volume" of contaminated soil was then assumed to be redistributed on the site surface, so that potential risks for a hypothetical future resident could also consider exposure to this soil (e.g., via incidental ingestion).

Location-Specific Analyses

We found that conducting location-specific analyses increased our flexibility to address site-specific conditions (including hot spots) and to iterate easily through exposure calculations when intake parameters were changed. For soil, a site-wide average derived from statistical analyses was used to estimate exposures and risks associated with random access to the site in the baseline risk assessment (i.e., the time spent was assumed to be evenly distributed across the site). However, this approach would masks risks from exposures at possibly "preferred" areas that might be more highly contaminated than others (e.g., a dump area rather than a non-source area). Therefore, risks were also estimated on a borehole-by-borehole basis for site soil. This approach addresses the fact that contaminants coexist differently in different locations across the site, and it permits a consideration of reasonable maximum exposures across multiple pathways and estimated risk distributions. For groundwater, the well-by-well analysis permitted the separate consideration of depths or locations for which the data were different for site-specific hydrogeological reasons. Risk isopleths were drawn from this analysis to support the determination of combined risks across multiple pathways by location.

STATISTICAL ISSUES AND FUTURE DIRECTIONS

The practical data-base issues discussed above address some of the challenges facing assessors of multimedia risks at Superfund sites. In the course of our analyses, we were able to develop or recommend a number of new or more comprehensive statistical approaches that could be used to enhance the methods applied for similar site-specific exposure and risk assessments. These approaches address the following issues: (1) treatment of measurements reported as below the analytical detection limits, (2) analysis of the spatial distribution of risks, (3) multivariate analysis and display of the spatial patterns of exposures and risks, and (4) assessment and presentation of uncertainties.

Measurements Below the Analytical Detection Limit

Chemical characterization data often include measurements that are reported as less than an established analytical detection limit (DL) -- i.e., non-detects. For the purpose of a risk assessment, it is useful to determine the true average of the population from which the sample was drawn in order to estimate values for the non-detects. If all samples below the DL are neglected in the averaging process, the resulting average will be an upwardly biased estimate of the true average. In fact, most methods for filling in values for samples below the DL result in biased estimates of the true average. Clearly, it is important to develop methods to calculate an unbiased estimate of the sample average when a portion of the data are below the DL.

Various researchers have recommended a number of statistical methods for estimating the mean of a series of observations that contain values reported as non-detects.[5-7] Two of the more sophisticated techniques, the maximum likelihood estimator (MLE) and the restricted maximum likelihood estimator (RMLE) methods, have been reviewed by Haas and Scheff[5]. Cohen[8] presented the MLE in terms of a tabulated function of two arguments; however, these tables are difficult to use in computer-based calculations. Haas and Scheff[5] developed a power-series expansion of the function that fit the tabulated values. The RMLE method developed by Persson and Rootzen[9] used two sets of estimators as approximations to the MLE, which are somewhat simpler to compute but do not appear to have been generally considered for the analysis of environmental data.

We have examined both the MLE and RMLE techniques by using soil chemical data for metals collected from the Superfund site studied. Figures 1 and 2 present histograms of the distributions of arsenic and aluminum in surface soil at the site. These two metals were selected because they exhibited different concentration profiles and were also detected in most of the soil samples. Using these data, we selectively censored the lowest 10% to 70% of the observations to (1) examine the influence of increasing the value of the DL in the estimation process and (2) test the reliability of the MLE and RMLE methods for predicting the average value for a censored data set. Neither technique provides a very good estimate of the mean without log-transforming the data. For arsenic, the MLE method applied to log-transformed data provided a better estimate of the mean concentration than did the RMLE method (i.e. the relative bias is much lower) for the entire range of censoring (Figure 3). The MLE technique also performs well for the censored, log-transformed aluminum data, and slightly better than the RMLE technique, with a relative bias of only a few percent (Figures 4 and 5). This finding, with the recommendation for log-transforming the original data prior to estimating the mean for a censored data set, is an important outcome of our assessment effort. In a further investigation of this technique (to be reported elsewhere), we also generated semi-empirical estimates for relative bias, depending on sample size and various measures of dispersion in the data such as geometric standard deviation and the coefficient of variation. We feel that the MLE method applied to log-transformed data is a superior method for estimating the mean of values that are reported to be less than the DL. This technique appears to be suitable for dealing with most, if not all, heterogeneous soil sampling data for which up to 50% of the values are below the DL.

Analysis of the Spatial Distribution of Risks

Exposures and risks associated with soil at this Superfund site were estimated three different ways: (1) quantifying the risks for the reasonable maximum exposure (RME) scenarios using site-wide data; (2) dividing the site into six separate areas on the basis of past use and contaminant information, then quantifying area-wide risks on a RME basis; and (3) quantifying risks for each of the individual (over 300) sampling locations. Each of these calculations were used to estimate chemical and radiological risks for a site worker and a hypothetical trespasser under current conditions and to a hypothetical recreational visitor to the site in the future, at which time it is assumed that the fence has disappeared and access is no longer restricted. (The site is currently fenced and public access is restricted by fences, locked gates, and security personnel.)

The upper bound health risks estimated on the basis of the RME assumption for a trespasser's exposure to chemicals in surface soil are summarized in Table I. These estimates assume exposures to C_{95} concentrations and are estimated on the basis of both site-wide analyses and location- and pathway-specific analyses. For each exposure pathway (ingestion, inhalation, and dermal), the predicted intake or dose corresponding to the RME exposure for each chemical was multiplied by the associated slope factor to derive a carcinogenic risk. The intake or dose was also compared to chemical-specific reference doses (RfDs) to derive a hazard quotient from which to assess the potential for noncarcinogenic effects. The chemical-specific cancer risks and hazard quotients were then summed across the various chemicals and appropriate pathways to estimate combined (total) carcinogenic risks and hazard indices for various receptors.

As shown in Table I, the variability of the RME risks between the six areas is small. Interestingly, the risks estimated for the six on-site areas are similar to those estimated for the off-site area. Risks from the dermal pathway (for which considerable uncertainty exists in the toxicological data) appear to dominate the estimated RME risks. Not surprisingly, predicted total RME risks using the site-wide C_{95} values were found to be similar to risks calculated individually for each of the six areas. Although we did not discern a clear spatial pattern in the distribution of estimated health risks for this Superfund site, nevertheless, we believe that the basic approach of grouping associated environmental data and risks could be useful in evaluating risks and focusing cleanup decisions for other sites (e.g., those with less widespread contaminant heterogeneity).

A more in-depth and focused assessment was accomplished by predicting risks on a borehole-by-borehole basis. The radiological and chemical cancer risks estimated for a current worker and for a hypothetical trespasser are shown in Figures 6 and 7, respectively. These figures indicate a fairly scattered distribution of incremental lifetime cancer risks in the range of 10^{-4} to 10^{-5} for the worker and 10^{-5} to 10^{-6} for the trespasser scenario. Predicted cancer risks greater than 10^{-4} are observed primarily in areas at which uranium and thorium disposal and spills, with subsequent runoff, are thought to have occurred. There is no discernible spatial pattern to these predicted "higher-risk" areas that are being targeted for cleanup considerations.

One of the limitations of EPA's guidance for the Superfund health risk evaluation is the recommended single-number RME approach. Ideally, it is desirable to display a full distribution of predicted risks from which the most appropriate strategy for risk management can be determined for a given site. The RME approach singles out an exposure point near the upper end of the possible exposure curve. The estimated risks corresponding to RME exposures are typically expected to be near the 95th percent upper confidence limit (UC_{95}) of the predicted risk distribution. Depending on the skewness of the contaminant concentrations, the correspondence between the RME (or C_{95}) risks and the UC_{95} risks based on the predicted borehole-specific risks may be weak.

We have investigated this problem by plotting the cumulative risk distribution functions of each exposure pathway for the predicted total hazard index and the predicted total cancer risk (Figures 8 and 9). These plots show that the risk distributions are not very skewed, even though the risks corresponding to the RME method are somewhat sensitive to the spatial structure of the data and to the compound-specific toxicity values. For the hazard index, the predicted RME value lies between the 80^{th} and 90^{th} percentile of the distribution of borehole- and pathway-specific risks. For cancer risk, the predicted RME value corresponds to a risk value approximately between the 80^{th} and 97^{th} percentile of the borehole- and pathway-specific risks. In each case RME-based risks generally seem to represent risks near the UC_{95} of the full distribution of predicted risks. Because the spatial variations in exposure and risk values were small for our soil analyses, the estimated C_{95} (RME) risks were quite comparable to the estimated UC_{95} risks. However, in other situations where the environmental data are much more skewed, these observations might change (see Fingleton et al.[10], a companion paper presented at this conference [Paper 91-172.1]).

In conclusion, we strongly recommend characterizing the full exposure and risk distributions for each Superfund site in addition to using the RME approach to estimate risks on the basis of C_{95} values. For sound site cleanup and risk management decisions, a comprehensive quantification and display of these risk predictions can be very helpful.

Multivariate Analysis and Display of Spatial Patterns of Exposures and Risks

In our investigation, the results of the location-specific analyses were used to identify areas that could be targeted for cleanup in order to support estimates of the volumes of contaminated material that may be managed and the related cost-benefit analyses. Site topography and past use of certain areas were also considered during this boundary definition effort. Relative to future directions for this type of assessment, factor analysis techniques could be used to refine the boundaries of areas targeted for cleanup considerations on the basis of risks. Principal component analysis, cluster analysis and LISTREL models could be used to identify and map the underlying relationships and patterns that may exist among the contaminants. These multivariate analyses could be used to reveal the locations of distinct waste types and areas with the highest potential for adverse health effects. Spatial analysis and geographical information system (GIS) approaches could also be used with sophisticated statistical methods such as co-krieging and surface-fitting algorithms (e.g., the Cressman algorithm, Barnes/Achtemeier algorithm, and various response surface methods) to address multimedia and multi-contaminant data. Spatial display techniques could also be applied to a presentation of the distribution of uncertainties surrounding the RME risk predictions.

Assessment and Presentation of Uncertainties

A major area of concern for multimedia risk assessments is how to best develop conservative but reasonable exposure scenarios. Thus, it is important to quantify the nature and magnitude of uncertainties surrounding each of these estimates, which propagate through the entire calculation. Resulting multidimensional predictions will benefit from sophisticated multivariate and GIS-based display techniques used to convey information on the accuracy and precision of these predictions. The tendency is often to choose a single value or a set of values from a conservative risk assessment in order to err on the side of caution. Unfortunately, this approach may at times lead not only to miscommunication of the level of true risks but to a misallocation or overestimation of resources needed for site cleanup as well. A comprehensive analysis that relies upon state-of-the-art methods for dealing with data heterogeneities is clearly important to the sound assessment of multimedia risks.

ACKNOWLEDGEMENTS

Much of this work was supported by the U.S. Department of Energy, under contract W-31-109-Eng-38. The authors wish to recognize the data interpretation, verification, and validation efforts of Ken Meyer, Julie Herron, John Coniglio, and Marc Nelson.

REFERENCES

1. U.S. Environmental Protection Agency. "National oil and hazardous substances pollution contingency plan; final rule," Federal Register, 55:8666 (1990).

2. U.S. Environmental Protection Agency. Risk Assessment Guidance for Superfund, Volume I, Human Health Evaluation Manual (Part A), EPA/540/1-89/002, Office of Emergency and Remedial Response, Washington, DC, Dec. (1989).

3. U.S. Department of Energy. "Radiation protection of the public and the environment," Order 5400.5 (1990).

4. U.S. Environmental Protection Agency. "Health and environmental protection standards for uranium and thorium mill tailings," 40 Code of Federal Regulations 192 (1990).

5. C.N. Haas, P.A. Scheff, "Estimation of averages in truncated samples," Environ. Sci. and Technol. 24:912 (1990).

6. C.C. Travis, M.L. Land, "Estimating the mean of data sets with nondeductible values," Environ. Sci. and Technol. 24:961 (1990).

7. M.A. Cohen, P.B. Ryan, "Observations less than the analytical limit of detection: a new approach," J. Air Poll. Contr. Assoc. 39:328 (1989).

8. A.C. Cohen, Technometrics 3:535 (1961).

9. T. Persson, H. Rootzen, Biometrika 64:123 (1977).

10. D.J. Fingleton, et al., "Stochastic model for estimating personal exposures in contaminated buildings at Superfund sites," Paper No. 91-172.1, presented at the 84th Annual Meeting of the Air and Waste Management Association, Vancouver, British Columbia (1991).

Table I. Estimated upper bound health risks associated with exposure to chemicals in surface soil (trespasser scenario)*.

Area	Total Hazard Index			Total Cancer Risk		
	Ingestion	Inhalation	Dermal	Ingestion	Inhalation	Dermal
area 0**	3.87E-04	4.49E-04	3.12E-02	1.11E-08	2.08E-09	7.47E-07
area 1	2.81E-04	4.45E-04	2.10E-02	2.47E-08	2.48E-09	2.36E-06
area 2	2.50E-04	4.54E-04	1.77E-02	1.49E-08	2.08E-09	1.09E-06
area 3	3.09E-04	4.08E-04	1.86E-02	2.77E-08	1.99E-09	1.99E-06
area 4	2.65E-04	3.96E-04	1.95E-02	1.54E-08	1.97E-09	1.09E-06
area 5	2.25E-04	4.68E-04	1.70E-02	1.47E-08	2.09E-09	1.14E-06
area 6	2.47E-04	3.87E-04	1.75E-02	1.87E-08	2.19E-09	1.20E-06
Total On-site Area	2.86E-04	3.94E-04	1.72E-02	2.29E-08	1.86E-09	1.46E-06

* Risks are for reasonable maximum exposure based on the estimated 95th percent upper confidence limit of the arithmetic mean concentration of each chemical.

** Off-site background near the site.

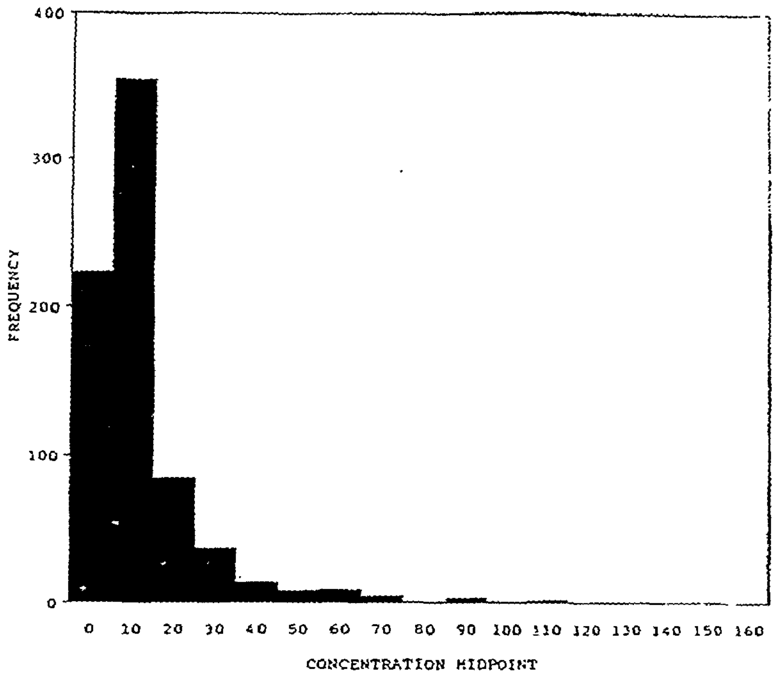


Figure 1. Histogram of arsenic concentrations (μg/g) measure in soil samples.

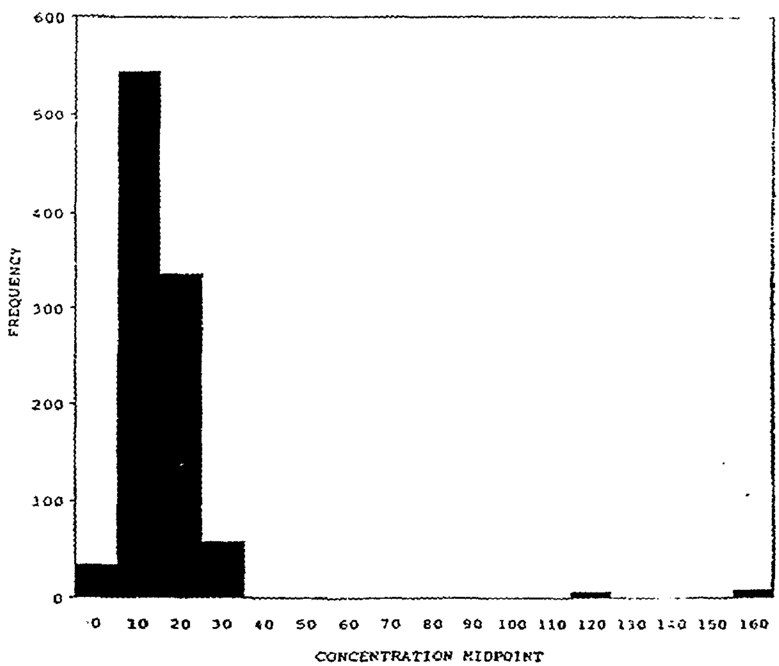


Figure 2. Histogram of aluminum concentrations (mg/g) measure in soil samples.

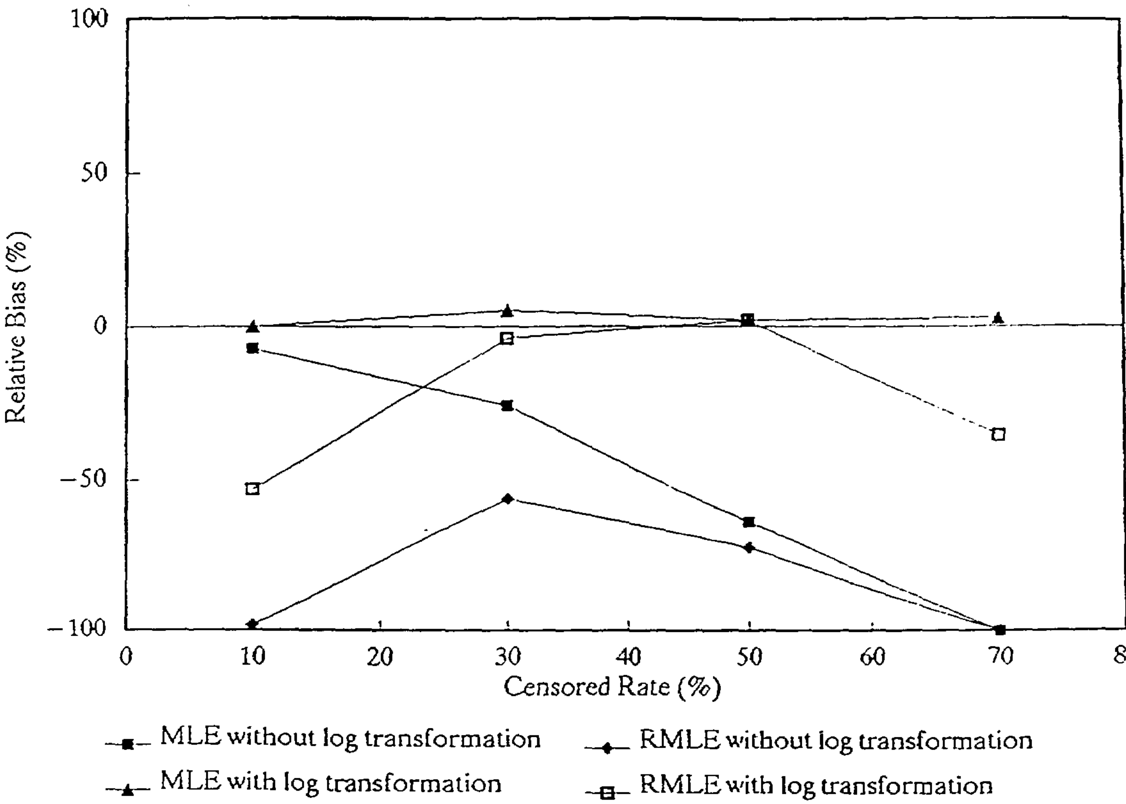


Figure 3. Relative bias of mean estimation for MLE and RMLE (censored data derived from arsenic concentrations measured in soil samples: N=714, mean=12.54, std=15.28).

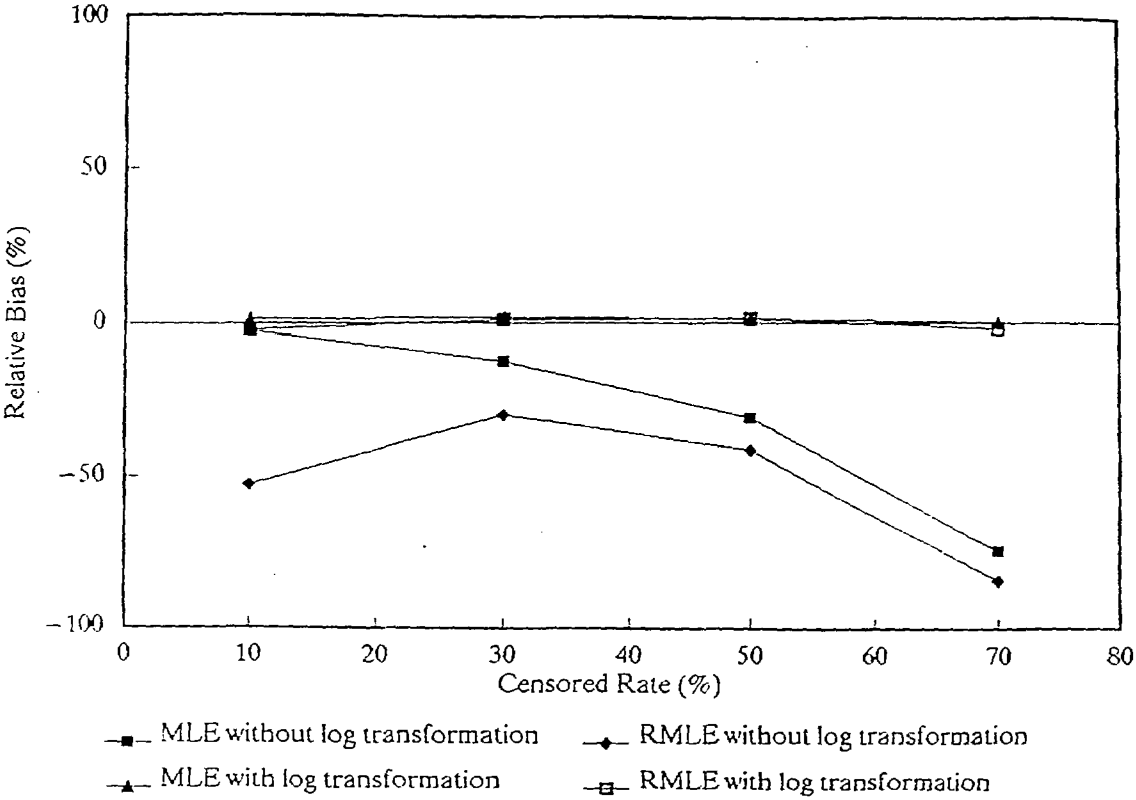


Figure 4. Relative bias of mean estimation for MLE and RMLE (censored data derived from aluminum concentrations measured in soil samples: N=956, mean=15,190, std=10,592).

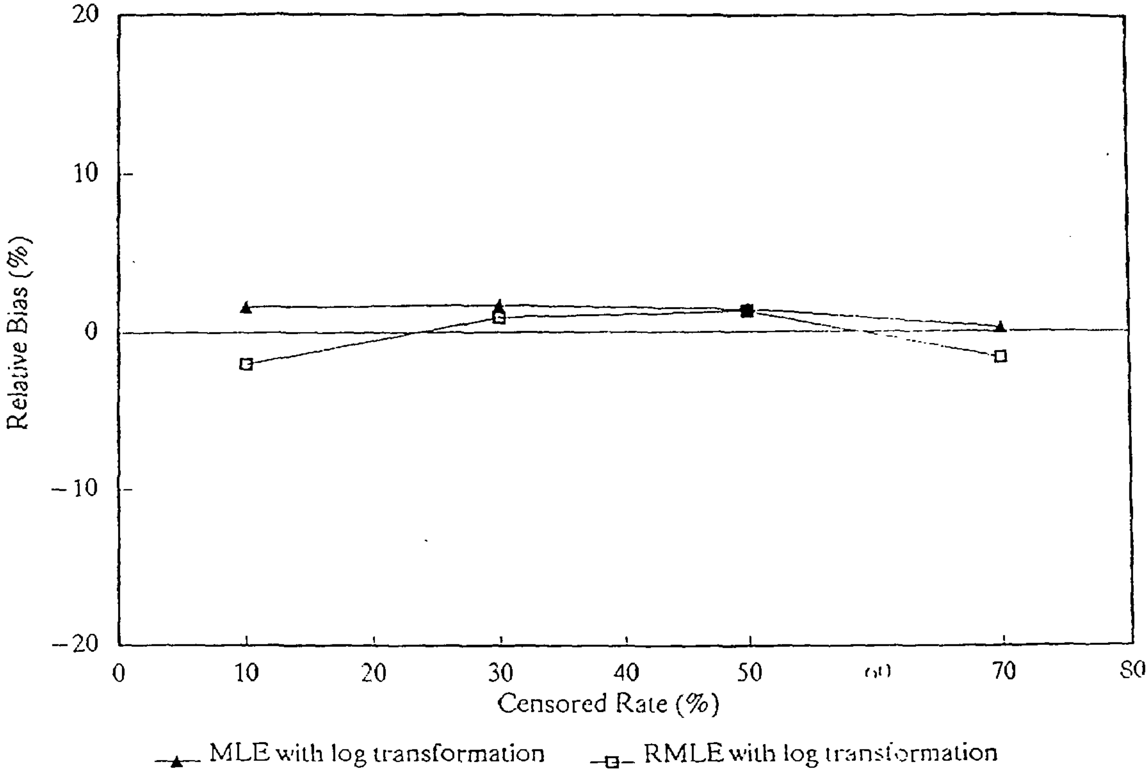
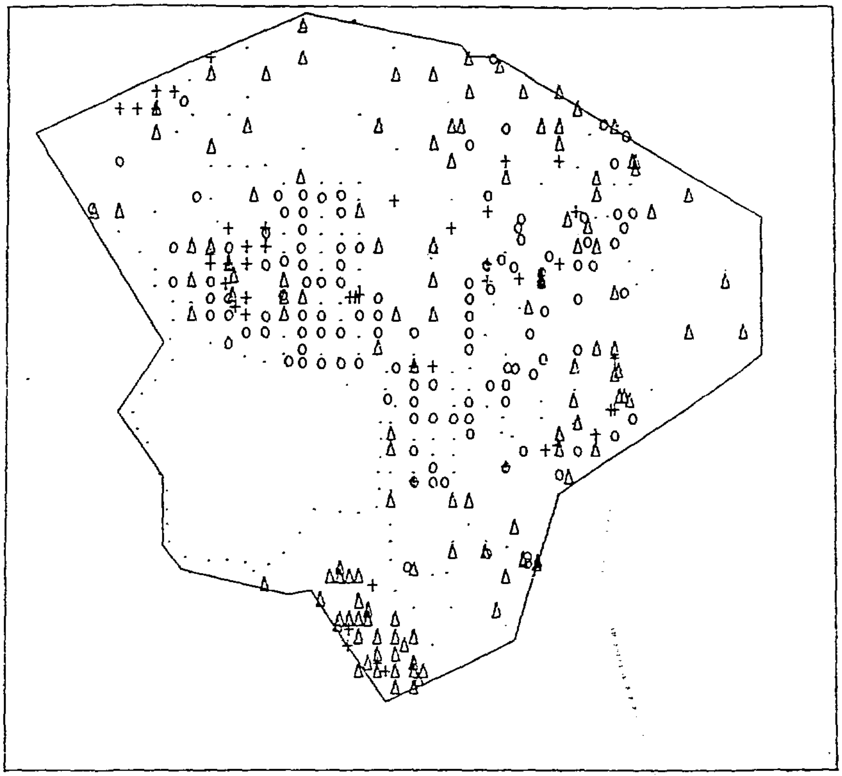
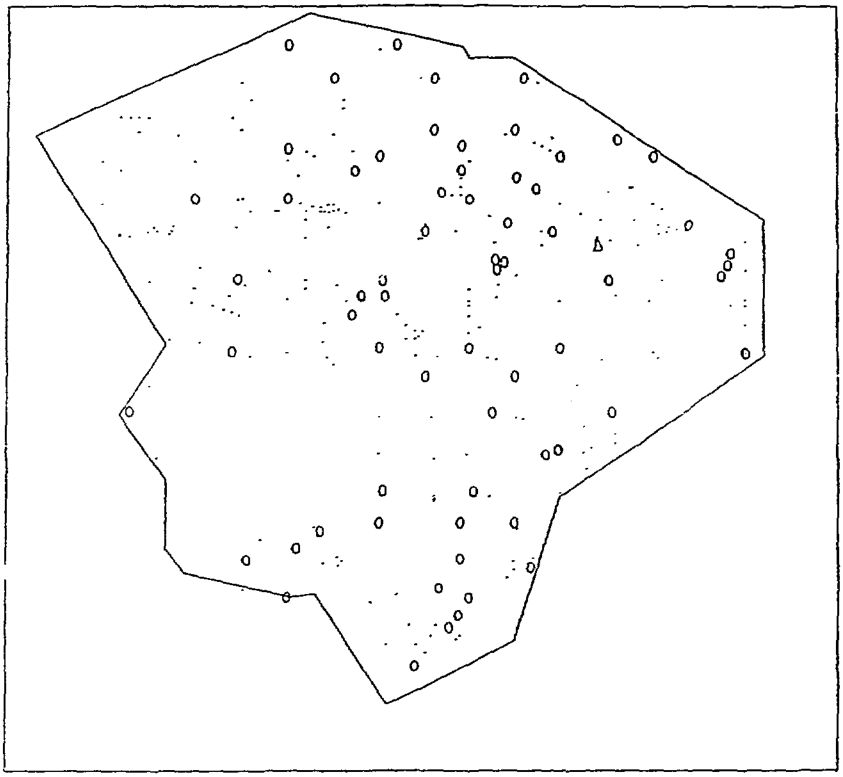


Figure 5. Relative bias of mean estimation for MLE and RMLE with log transformation (censored data derived from aluminum concentrations measured in soil samples: N=956, mean=15,190, std=10592).



Key:
- \+ risk greater than 1E–04
- △ risk between 1E–04 and 1E–05
- o risk between 1E–05 and 1E–06
- . risk less than 1E–06

Figure 6. Estimated cancer risks associated with cumulative exposure to radionuclides in surface soil (current worker scenario).



Key:
- \+ risk greater than 1E-04
- Δ risk between 1E-04 and 1E-05
- o risk between 1E-05 and 1E-06
- . risk less than 1E-06

Figure 7. Estimated cancer risks associated with cumulative exposure to chemicals in surface soil (trespasser scenario).

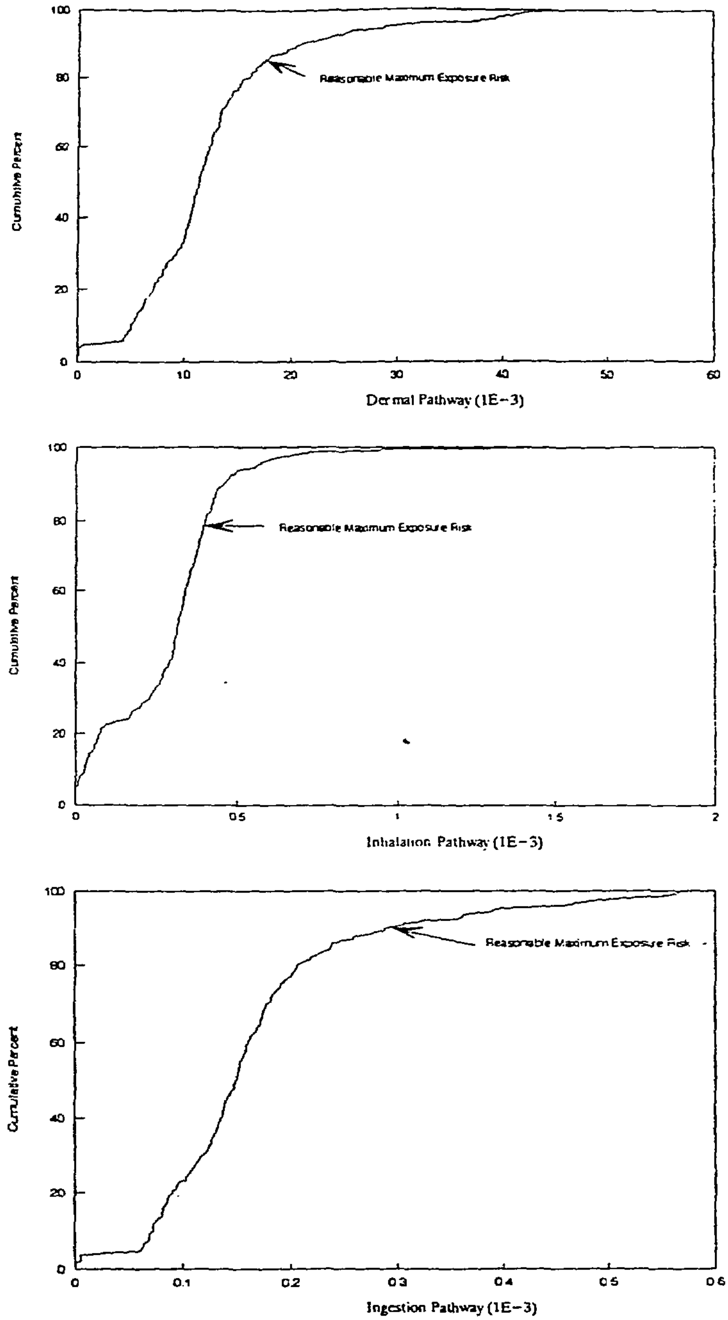


Figure 8. Cumulative distribution of predicted total hazard indices associated with pathway-specific exposures to chemicals in soil (trespasser scenario).

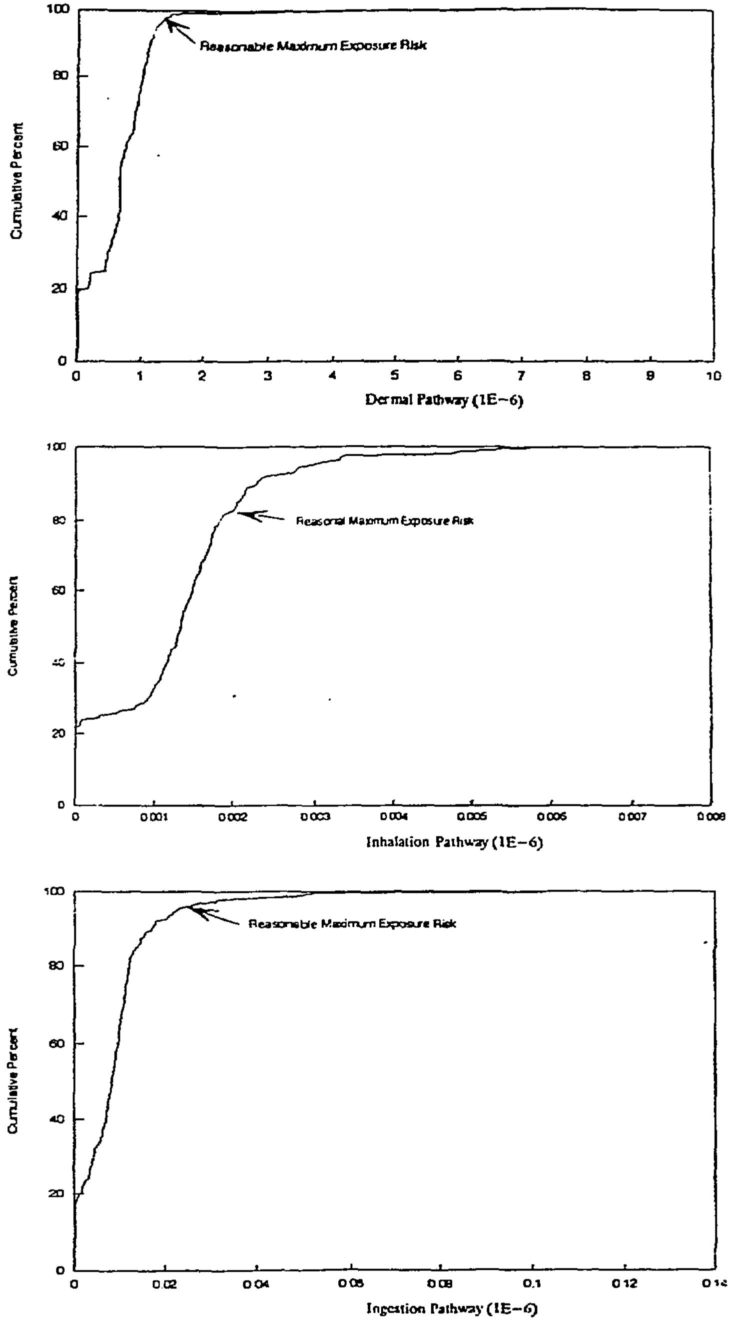


Figure 9. Cumulative distribution of predicted total cancer risks associated with pathway-specific exposures to chemicals in soil (trespasser scenario).