US EPA ARCHIVE DOCUMENT

3.0 Overview of the 3MRA Science Methodology, Data Sets, Uncertainty, and Implementation

US EPA ARCHIVE DOCUMENT

The genesis of the underlying science of 3MRA Version 1.0 followed from work of EPA in the late 1980's, in development of the concept of toxicity characterization for defining separate approaches for management of hazardous and nonhazardous solid waste. A brief review is provided here on key concepts of the historical perspective and science underpinning the 3MRA technology, as well as data collection approaches for the national assessment, with a particular focus on related aspects of uncertainty analysis (UA) and sensitivity analysis (SA).

In response to the 1995 SAB HWIR review, OSW and ORD initiated a collaboration to define an improved science approach to adequately address the SAB's associated review comments and recommendations. The outcome of this reflection was later captured in a joint OSW-ORD science-planning document for development of the 3MRA software technology. This 3MRA science-planning document was extensively peer-reviewed, and resulted later in a formal science methodology document entitled *Framework for Multimedia, Multipathway, and Multireceptor Risk Assessment (3MRA)* (Marin *et al.*, 1999; see Appendix A). The conceptual layout of the underlying science approach was also recently published by Marin *et al.* (2003).

To capture the detailed technical aspects of the planned risk assessment methodology, the reader is referred to the science plan and original methodology document (Appendix A). Key aspects of the technological implementation of the plan will be summarized in Section 4 and 6 of this Volume IV. The reason for its presentation here is to underscore the importance of this science planning effort, drawing upon key issues already discussed in Section 2, and to qualify various elements of the UA/SA plan for 3MRA delineated in Section 9 of this Volume IV.

3.1 Historical Perspective of Hazardous Waste Management and HWIR

In capturing salient points of the science methodology with respect to UA/SA, most of the historical materials cited in the following materials were excerpted directly from Marin *et al.*, (1999), with minor edits, and, as such, are by-and-large attributed to these authors. The reader is referred to the original document relative to the immediate discussion that follows.

The U.S. EPA originally developed a methodology to set regulatory threshold levels for chemical constituents in wastes, based on the expected groundwater impact of these constituents leaching from Subtitle D waste management units (WMUs). In 1990, the Agency applied the methodology to develop the final toxicity characteristic (TC) approach still used today to identify

hazardous wastes. According to the TC approach, a solid waste is deemed hazardous based on its potential to leach significant concentrations of specific toxic constituents. In this approach, if a waste with constituents exceeds any of the corresponding Toxicity Characteristic regulatory limits, the waste is classified as toxic hazardous waste.

In December of 1995, the Agency proposed to amend existing regulations for disposal of listed hazardous wastes under RCRA. The 1995 proposal (60 FR 66344, December 21, 1995) outlined the Hazardous Waste Identification Rule (HWIR, 1995), which was designed to establish constituent-specific exemption levels for low risk solid wastes. Wastes applicable for consideration under HWIR were those designated as hazardous because they were listed, or had been mixed with, derived from, or contained the listed wastes. Under the HWIR proposal, waste generators of listed wastes that could meet the new, concentration-based exemption level criteria defined by the HWIR methodology would no longer be subject to the hazardous waste management system rules, specified under Subtitle C of RCRA for those wastes. Basically, HWIR would establish a risk-based "floor" for low risk hazardous wastes that would encourage waste minimization, and the development of innovative waste treatment technologies.

The "mixture" rule and the "derived-from" rule were promulgated as part of the first comprehensive regulatory program for the management of hazardous wastes under RCRA in May of 1980. The mixture rule defined as a hazardous waste any solid waste that is mixed with one or more listed wastes, and the derived-from rule labeled as hazardous waste any solid waste generated from the treatment, storage, or disposal of a listed hazardous waste. Both have been considered important definitions in regulating the disposal of hazardous wastes consistent with reducing risk to human health and the environment. However, since these two rules apply regardless of the concentration or mobility of the hazardous constituents associated with the solid wastes, the potential for over-regulation is a possibility. One of the primary purposes of HWIR was to provide a risk-based methodology for identifying possible instances of over-regulation, and to provide an avenue for relief from the Subtitle C disposal regulations, as appropriate.

In the HWIR95 framework (HWIR, 1995), groundwater and non-groundwater pathways were separately analyzed. A review by the EPA's Science Advisory Board indicated that in using this approach it would be difficult to maintain mass balance, and may lead to significant, but unknown, errors in the exposure estimates. How far the results would diverge from those of a true multipathway approach could not be determined without going through a number of representative multipathway calculations. The SAB recommended that the non-groundwater pathway framework used for HWIR95 be abandoned in favor of true multipathway calculations (EPA, 1996b). In response to the SAB's recommendations, a consistent multimedia, multireceptor, multipathway risk assessment (3MRA) approach was conceptually formulated. In addition, a conceptual risk assessment procedure for determining safe waste concentration limits for hazardous chemicals and metals using a finite-source 3MRA approach was outlined. The theoretical approach included a Monte Carlo algorithm that would facilitate the calculation of the uncertainty in the proposed regulatory waste concentration levels as a function of input parameter sampling and measurement errors, and prediction model errors.

US EPA ARCHIVE DOCUMENT

Implementation of this Monte Carlo-based assessment strategy could conceptually be fully implemented, provided that all of the required model components, input data, and necessary computational resources were available. However, the scope of its actual implementation would, of course, depend on the availability of data, computational resources, and time constraints. In any event, the outcome of the 1995 SAB review placed significant demand for a more formally developed technology (i.e. 3MRA Version 1.0) to assess associated national risk from land-based disposal of hazardous contaminants, along with commensurate capabilities to evaluate uncertainty and sensitivity of 3MRA modeling system predictions.

3.2 3MRA Science Plan Objectives

Already summarized in the problem statement presented in Section 1.2, the objective of the science plan was to outline a conceptual framework for the finite source, Multimedia, Multipathway, Multireceptor Risk Assessment (3MRA) methodology. As originally conceived by OSW, the proposed methodology was not meant to be definitive, but rather was intended to provide a foundation for the development of a modeling system technology that would facilitate a technical and quantitative response to the problem statement for national-scale regulatory decision-making. For example, there were a number of options available for the regulatory objectives, including issues associated with the definition of protection measures, definition of the problem statement, and the number and type of receptors to be evaluated, to name just a few (Marin *et al.*, 1999). The proposed framework was intended to be sufficiently general to accommodate alternative options.

Perspective on the methodology's underlying foundation, objectives, and analytical basis is summarized from Marin *et al.* (1999) by the following statements:

National-Scale Risk Assessment Approach

- The 3MRA methodology will be risk-based, where the constituent-specific waste exemption levels (i.e., exit levels for "listed" hazardous waste) will be set such that no significant risk to human or ecological health will occur as a result of the disposal of the waste in non-hazardous waste management units.

- The exit levels will apply to all waste streams under all applicable Subtitle D land-based WMUs scenarios including landfills, wastepiles, land application units, surface impoundments, and will be extended as well to aerated tanks.

- The assessment of potential health risks will be conducted separately for both human and ecological receptors, and exit levels will be based on the most limiting concern.

- The assessment will be a screening-level, risk-based evaluation of potential risks resulting from long-term (chronic) exposure to chemicals and metals released from such WMUs.

US EPA ARCHIVE DOCUMENT

- The assessment will be national in scale and site-based, that is, risks will be assessed at individual sites across the U.S. where WMUs may be located, where the resulting national distribution of risks will form the basis for establishing exemption criteria.

- For each site, statistically sampled from a national database of WMUs, the simultaneous release of chemicals or metals from the WMU to each environmental medium, the fate and transport of the chemical or metal through a multimedia environment, and the receptor-specific exposures and risks will be simulated.

- An exit level for each WMU type will be derived independently for each chemical. The effects of the different WMU types within a site will be considered separately in the exit-level determination decision context.

- The end point of the technical assessment will be a compilation of the risks to form a national scale joint distribution, reflecting the relationship between chemical concentration in waste streams and human and ecological health risk.

- Specific exemption levels will be selected from these distributions on the basis of Agency policy (e.g., appropriate degrees of protectiveness, receptor types, sites, distance from units, geographic location).

- The resulting chemical-specific exit levels represent threshold waste concentrations below which the associated wastestream is not considered hazardous and therefore does not require Subtitle C type disposal.

Fate and Transport

- A given site will be defined by the area contained within a 2 km distance from the WMU boundary, as defined by the WMU area.

- The evaluation of the impact on receptors of concern will be performed for a fixed time, beginning at time t_0 until T_{max}. The value of T_{max} varies for different chemicals and metals, but will not exceed 10,000 years as set by Agency policy.

- The total mass of a given chemical or metal to be managed in a WMU will be a finite value.

- Mass balance in the waste management unit will be maintained at all times. If the mass in the unit is exhausted through releases to the environment and/or degradation, no additional releases can occur from the unit.

- The fate/transport components take the source releases from a WMU and distribute the mass through each medium to determine the concentrations of the chemical for each contact medium (e.g., air, groundwater, soil, surface water, plants), in each exposure area, from time t_0 to T_{max}.

US EPA ARCHIVE DOCUMENT

- The contaminant concentration for any contact medium, at any point within an exposure area at a given time, will be given by the areal average over that area.

Exposure Pathways and Receptors

- Human receptors include child and adult; residents, home gardeners, beef and dairy farmers, and recreational fishers. Exposure pathways include inhalation of outdoor air and shower air and ingestion of contaminated drinking water, garden and farm products and fish.

- Ecological exposure and risk will focus on individual effects related to population and community viability within habitats found in the proximity of sites.

- The assessment will include an estimation of the potential exposures per exposure pathway/receptor, and aggregated across pathways, followed by an estimate of the resulting carcinogenic (humans only) and noncarcinogenic health effects.

- The impacts to receptors of concern are evaluated for each chemical or metal independent of the effects of other chemicals and metals. The cumulative effects of different chemicals, acting simultaneously on a receptor is not considered.

- Each site encompasses one or more exposure areas (or sectors). Receptors of each type in each exposure area are represented by a single receptor (representative receptor) with a weight corresponding to the total number of receptors of that type in that exposure area.

- A receptor may be exposed simultaneously via multiple pathways, each involving different combinations of contact media and exposure routes.

- Human exposure routes to be considered include inhalation and ingestion. The dermal contact route will not be considered because of limited data. Exposure media for human receptors include groundwater, soil, air, and biota.

- Ecological exposure routes include ingestion and direct contact. Exposure media for ecological receptors include surface water, soil, and biota.

- Each receptor type in an exposure area at a site will be represented by a series of T_{max} longitudinal cohorts. Each longitudinal cohort corresponding to a given receptor type has identical exposure characteristics with the exception that the initial exposure conditions for successive cohorts are lagged by one-year interval from time t_0 to T_{max}.

3.3 Science Based Modeling Approach

Essentially, the 3MRA national assessment modeling strategy led to a technology design (3MRA Version 1.0) that uses a regionalized, site-based approach to estimate risks from waste

US EPA ARCHIVE DOCUMENT

management units on a regional or national basis. This section presents a brief overview touching on elements of the technology developed to implement the science-based modeling approach outlined in Section 3.2 and described fully in Appendix A.

3.3.1 Science-Based Models

Providing for solution of the problem statement in context of the science-based risk assessment methodology, 17 separate science-based models (i.e., modules of the 3MRA modeling system) were eventually developed or acquired as legacy codes. These modules, discussed in detail in Volume I and summarized in Section 4, cover the following key areas:

- **Contaminant Source-Term (WMU) Models**
    a. Landfill
    b. Wastepile
    c. Aerated Tank
    d. Surface Impoundment
    e. Land Application Unit

- **Fate and Transport Models**
    a. Airshed
    b. Watershed
    c. Vadose Zone
    d. Aquifer
    e. Surface Water

- **Foodweb Models**
    a. Aquatic
    b. Terrestrial
    c. Farmfood

- **Exposure Models**
    a. Ecological
    b. Human

- **Risk Models**
    a. Ecological
    b. Human

3.3.2 Runtime System Level Processors

In addition, several system level processors would be needed to organize and integrate the risk assessment methodology and supporting data into a useable technology. The primary processors used at modeling system runtime include:

- **System User Interface (SUI)**
    - An interface for the user to run the modeling system

US EPA ARCHIVE DOCUMENT

- **Site Definition Processor (SDP)**
  - Extracts data from 3MRA databases for individual model runs

- **Chemical Properties Processor (CPP)**
  - Provides chemical data needed for individual model runs

- **Multimedia Multipathway Simulation Processor (MMSP)**
  - Manages execution of the science modules for a given model run

- **Exit Level Processor I (ELP1)**
  - A data compression and risk protection summary output processor launched per model run

- **Exit Level Processor II (ELP2)**
  - A data analysis tool that queries final risk summaries and calculates exit levels across model runs based on several user-defined decision variables.

3.3.3 Exit Level Calculation Steps

Touching further on the actual technological aspects developed to implement the 3MRA science methodology, the basic calculation steps of the science-based risk assessment approach are outlined below, and are further described in Section 4.6. The calculation flow is laid out in graphical form in Figure 3-1:

- Step 1 → Select Facility/Environmental Setting from Representative Sample
- Step 2 → Select WMU/Source Type (and effectively waste type; e.g., liquid)
- Step 3 → Select Chemical Constituent
- Step 4 → Select Concentration of the Constituent in Waste, C_w
- Step 5 → Read Data Files
- Step 6 → Select Receptor Location or Habitat
- Step 7 → Calculate Contact Media Concentration at Receptor Location
- Step 8 → Select Receptors
- Step 9 → Select Receptor Age Cohort
- Step 10 → Select Pathway
- Step 11 → Calculate Risk or Hazard
- Step 12 → Select Next Pathway
- Step 13 → Select Next Age Cohort
- Step 14 → Select Next Receptor
- Step 15 → Select Next Receptor Location
- Step 16 → Find the Critical Time Period for Evaluation
- Step 17 → Output Risk/Hazard at T_{crit} (ELP1)
- Step 18 → Select Next C_w
- Step 19 → Select Next Chemical
- Step 20 → Select Next WMU

US EPA ARCHIVE DOCUMENT

- Steps for Uncertainty Analysis → Insert Inner and Outer Loops As Needed
- Step 21 → Calculate Exit Levels (ELP2)

Figure 3-1 illustrates how the 3MRA model conceptually executes a given scenario set defined by the user through the System User Interface (SUI). The "internal" model looping order (Steps 5 to 17 in Figure 3-1) is always the same, where the SUI issues individual deterministic model runs by looping first through waste stream concentration (C_w), then through chemicals/metals, then through random Monte Carlo iterations, next through sources, and finally through sites. Generically, the term chemical used in various figures throughout this discussion, infer the collection of 43 organic chemicals and metals evaluated in 3MRA, where notably, Mercury, as a single metal of concern, is represented by 3 separate species. Discussed in Section 2.6.6, various uncertainty and sensitivity analyses strategies (i.e., 2nd-order analysis, etc.) would be implemented through management of output data in steps via the insertion of inner and outer loop actions. The actual "internal" and "external" model looping order (Steps 1-4 and 18-21) would be immaterial to the ELP2 calculation outcome, depending on the predictive uncertainty analysis or sensitivity analysis simulation design.

3.4 Sources of Data and Uncertainties for the National Assessment Strategy

Described by Marin *et al.* (1999) and discussed in Section 2.6.2, the input parameters for the proposed framework are used to define the modeling scenario for a facility [i.e., a site] and can be grouped into four general classes:

- [3MRA Class 1 Inputs] - Variables that describe the characteristics of the waste management facility, including area and depth,

- [3MRA Class 2 Inputs] - Variables that describe the environmental conditions of the facility and its surroundings including hydrologic, hydrogeologic, meteorological, and geochemical conditions at the site,

- [3MRA Class 3 Inputs] - Variables that describe the (physiologic and behavioral) exposure and response characteristics of the receptors; and

- [3MRA Class 4 Inputs] - Variables that describe the physical, chemical, and biochemical properties of the chemical constituents.

Section 2.6.2 also describes various sources of uncertainty identified by Marin *et al.* (1999). Under ideal conditions, the 3MRA Monte Carlo approach to assess uncertainty in the national risk assessment strategy would be based on the following data (Marin *et al.*, 1999):

1. A statistically-designed sample of waste management units from the target population of waste management units (WMUs) in the U.S.,

2. Direct measurement of the facility/site characteristics (e.g., unit area and volume; depth to groundwater; aquifer thickness; hydraulic conductivity; hydraulic

US EPA ARCHIVE DOCUMENT

gradient; distance to nearest well; number, location and physiologic/behavioral characteristics of receptors) at each sampled site; and

3. Availability of calibration/validation data sets to estimate data measurement and component model prediction error structures.

Marin *et al.* (1999) originally conceived that the national scale assessment would rely as much as possible on a site-specific data collection and modeling approach for the source (other than waste characteristics), fate and transport, and exposure characteristics of a facility. In the absence of site-specific data, data from regional and national distributions would be used. The Industrial D Screening Survey (Westat, 1987) used contacted 15,844 industrial facilities. Of those, 2,850 reported that they managed waste in a landfill, LAU, surface impoundment, or waste pile. To adequately represent the 17 industry groups in the survey, EPA randomly selected 201 sites from these 2,850 facilities (with 419 site-WMU combinations). The basic assumption for the national risk assessment is that these data are representative of the types, characteristics, and locations of nonhazardous industrial WMUs. Regional assignments by site, including meteorological station, USGS hydrologic region, and groundwater class were then constructed.

Data collection efforts eventually led to the creation of 5 major databases used in 3MRA model runs:

- Site-based database (Class 1 and many Class 2 inputs)
- Regional database (Vadose and aquifer zone parameters, stream base flow rates)
- National database (Many Class 2 and Class 3 inputs)
- Meteorological database (Best available data from nearby meteorological stations)
- Chemical properties database (Class 4 inputs)

3.4.1 Sources of Data Actually Collected By Science Module

Further described in Volumes I and II and Section 4, the 3MRA modeling system contains 17 media-specific pollutant release, fate, transport, exposure, and risk modules. The modules have varying data requirements covering a wide range of general data categories:

- Waste management unit characteristics,
- Waste properties,
- Meteorological data,
- Surface water and watershed layout and characteristics,
- Soil (vadose zone) properties,
- Aquifer (saturated zone) properties,
- Food chain or food web characteristics,
- Human and ecological exposure factors,
- Types and locations of human receptors and ecological receptors and habitats surrounding a WMU, and
- Chemical-specific data including chemical properties, bio-uptake and bioaccumulation factors, and human and ecological health benchmarks.

US EPA ARCHIVE DOCUMENT

Summary information on sources of data for the above categories, taken directly from Volume II, is described in the Sections 3.4.2 through 3.4.13 along with excerpted highlights on various aspects of input data uncertainty. The categories are generally delineated in Table 3-1 regarding major database sources in 3MRA, inferring the various data collection approaches used (i.e., national, regional, and site-based or site-specific). Detailed descriptions of final sources of data and characteristics of individual model inputs (and outputs) are also described module-by-module in Section 8.

3.4.2 Waste Management Unit

Three parameterization approaches were used for WMU inputs: site-specific, site-based and national. Site-specific data on WMU area, capacity, and waste loading rates were obtained from the Industrial D Screening Survey (Westat, 1987). Size-related WMU variables, which were derived from these Industrial D data, are referred to as site-based data. All other WMU inputs were developed on a national basis as either distributions or fixed values, depending on potential variability and model sensitivity.

For the nationally collected data, one issue of concern was the availability of recent data on actual Industrial D units. EPA considered, however, the approach of basing the assessment on actual WMU, land use, and population data to be preferable to developing and evaluating hypothetical exposure scenarios. Another issue was whether to revisit the methodology used to screen out questionable entries in the Industrial D database. For consistency, it was decided to use the methodology from previous EPACMTP modeling efforts, as described in U.S. EPA (1997c). Finally, the tank database may under-represent highly aerated tanks. This under representation introduces some uncertainty into the analysis, the result of which is that risks from highly aerated tanks may be underestimated.

3.4.3 Waste Properties

The waste property variables should be reflective of nonhazardous industrial wastes. However, little data were available on the waste characteristics needed by the models. EPA conducted an investigation and review of available waste characterization data, but results are incomplete and were not adaptable in time for 3MRA model runs. Instead, engineering judgment was used to estimate reasonable ranges and typical variables for properties, and a uniform or triangular distribution was assumed to represent uncertainty.

A potential source of uncertainty, by assumption waste property data are not treated through correlation structures within 3MRA.

3.4.4 Meteorological Database

Meteorological data were collected regionally by meteorological station, with each of the 201 Industrial D sites included in the representative national data set assigned to the nearest station with similar weather conditions and adequate weather data for the analysis. In making these assignments, EPA considered all available data from 199 meteorological stations across the conterminous United States to find the best data for each site. This process resulted in 99

US EPA ARCHIVE DOCUMENT

meteorological stations being matched to the 201 industrial sites. Most meteorological data were extracted from Solar and Meteorological Surface Observation Network (SAMSON) hourly data files and aggregated as necessary to daily time series, monthly time series, annual time series, and long-term averages.

The meteorological data set collected represents the largest, most comprehensive set of meteorological data ever prepared for a risk assessment modeling effort. All readily available meteorological data were utilized for the effort to ensure that the most applicable data are used for each of the 201 Industrial D sites modeled. In addition, extensive quality control was conducted to ensure that the data were accurate and complete, and to identify and correct data gaps. The great volume of data (5 data files for 99 meteorological stations) required extensive automated data processing to compile and calculate the meteorological data required by the 3MRA modeling system. Most of the issues and uncertainties associated with this effort have to do with the assumptions and simplifications necessary to write the automation programs, or uncertainties associated with replacing missing data. Analyst bias was largely mitigated by the automated procedures applied.

3.4.5 Surface Water And Watershed Layout and Characteristics

Site-based, delineation of the 3MRA modeling system watershed/waterbody layouts at the 201 sites used both geographic information system (GIS) and conventional database programs. The GIS programs were used to: (1) compile the hydrologic, topographic, land use, and wetlands data coverages needed for delineating and attributing watersheds and waterbodies; (2) extract site-specific data from these data sets for each of the 201 Industrial D sites; (3) delineate watershed subbasins, waterbodies, and local watersheds; and (4) export the resulting spatial parameters in data tables for further processing. Overarching uncertainties include inconsistencies in scale for several data sources, the resolution of the digital elevation model (DEM) data with respect to the area of influence (AOI), data gaps from incomplete or inconsistent coverages for certain data, and site-to-site methodology differences necessitated by variability in both data and site characteristics.

Regional and national data were collected to supplement the site-based data set. These include regional water quality and flow data extracted from the EPA's Storage and Retrieval System (STORET) database and national ranges and distributions from literature and professional judgment. The primary issues and uncertainties for both regional and national surface water data arise from the lack of readily available site-specific data and are therefore associated with the representativeness of the regional and national data for a particular site. Like all other national and regional data sources described here, this issue is generally addressed in the national assessment uncertainty analysis (Section 2.6.3). For regional data, site-to-site variability and accuracy was preserved to the extent practicable by keeping the region over which data were collected and statistics compiled as small as possible. National data were only collected where site-based and regional data were not available.

US EPA ARCHIVE DOCUMENT

3.4.6 Soil (Vadose Zone) Properties

In terms of collection scale, all soil parameters can be considered to be site-based in that they are either site-specific or derived from site-specific data using national relationships. The latter category includes soil hydrologic properties derived from site-specific soil texture or hydrologic class and other properties derived from a combination of soil texture or class and site-specific land use. Site-specific soil data were collected for all 201 sites, largely using automated methods. Geographic information system (GIS) programs were used to identify and extract, for each Industrial D site modeled, soil map units and map unit areas by watershed subbasin and by waste management unit (WMU). Database programs were used to extract soil data from the underlying databases by these map units and process them to generate soil properties by watershed for surface soil, and by the WMU for the entire soil column (vadose zone or subsoil).

In general, the site-specific soil data readily available from STATSGO, and its associated databases, are considered more than adequate for a national screening analysis. Although soil properties do vary significantly on a much smaller spatial scale than the nationwide soil data in STATSGO, the 3MRA modeling system, which assumes average soil properties across a watershed or vadose zone, could not take advantage of such variability. Given the demonstrated quality of the soil data in STATSGO, the QA/QC measures designed to ensure effective data processing and transfer, and the national scale of the 3MRA modeling system, EPA does not believe soil data are a significant source of overall uncertainty in the national assessment. Minor uncertainties associated with the soil data include: (1) the assumption of homogeneity over map units; (2) in the few cases when data were missing, an infill value, representing the most common value nationally, was substituted; (3) the scale at which data were available for the local watershed; and (4) lack of data currency and the scale of the data associated with land uses.

3.4.7 Aquifer (Saturated Zone) Properties

For most aquifer variables, site-specific data were not available given limited project resources. Instead, EPA relied upon a regional site-based data collection approach and data very similar to that used in the risk assessment that supported the 1995 Hazardous Waste Identification Rule (HWIR, 1995). Some inconsistencies were noted in that review between the ground water data and data collected for the other pathways, and EPA has addressed these concerns. First, site-specific soils data provide a consistent set of soil properties to all 3MRA modeling system modules requiring such data, including the ground water pathway module. Second, receptor wells (private drinking water wells) are placed around each site using U.S. census data for all exposure pathways. Third, ground water flow direction is defined site-specifically based on topography and surface drainage patterns.

Four key hydrogeologic parameters were collected regionally based on a site's hydrogeologic environment: hydraulic conductivity and gradient, and saturated and unsaturated zone thickness. These parameters were characterized using discrete distributions derived from the American Petroleum Institute's (API's) Hydrogeologic Database (Newell *et al.*, 1989). Based on a site's regionally determined groundwater class assignment, the system randomly picks one set of the four variables for each Monte Carlo realization. By selecting the variables together, the dependencies between the hydrogeologic parameters (which represent measurements taken at

US EPA ARCHIVE DOCUMENT

the same location) are preserved in the analysis. The remaining aquifer variables are characterized by national distributions.

Aside from uncertainties quantified in the Monte Carlo Simulation design (2.6.3), no site-based ground water flow direction data were available, and ground water was assumed to flow in the same direction as surface water in the local watershed. In addition, the local watershed-to-reach connectivity was used to represent the aquifer-to-reach connectivity at the site. This adds uncertainty to the analysis, not explicitly addressed, because ground water flow does not always follow topography or discharge to the nearest waterbody.

3.4.8 Food Chain Or Food Web Characteristics

Data for the farm foodchain and terrestrial foodweb modules were collected to quantify parameters required to develop exposure profiles for receptors in these respective food webs. These parameters include bioconcentration factors, biotransfer factors, partitioning coefficients, and ingestion rates that the modules use to estimate movement of constituents through food webs. Many of the parameters required for the farm food chain are also applied to the terrestrial food web. The selected receptors and their respective prey preferences constitute the primary differences between the farm food chain and the terrestrial food web; however, many of the exposure relationships are similar. Two general databases were generated for these food webs: a chemical-specific database and an exposure-related database. The chemical-specific parameters vary depending on the constituent evaluated, whereas the exposure-related database consists of parameters that represent uptake and deposition rates, empirical correction factors, and adjustment fractions to account for reductions in potential exposures.

Two primary databases support the aquatic foodweb module: the aquatic food web chemical properties database and the fish attribute database. The associated chemical properties database contains two types of parameters: (1) equation variables used in estimating the bioconcentration and bioaccumulation of nonionic organic compounds into aquatic organisms using chemical-specific properties; and (2) experimentally derived bioconcentration and bioaccumulation factors (BCFs/BAFs) for ionic compounds, such as metals. The fish attribute database is composed of data characterizing the physiological traits and dietary preferences of aquatic biota. This database characterizes the life history attributes that influence the exposures of fish and other aquatic food web biota. Life history parameters such as fish body weight, tissue lipid fraction, tissue water fraction, and common prey items are identified. This database also identifies the types of fish that are likely to be consumed by humans.

Except for national-based water consumption rates for beef and dairy (triangular distributions), all input parameters, including those used to parameterize functional relationships, are currently defined on a national basis by constant values, where some data was compiled from regional data sources. Data collection was generally constrained by the lack of available information needed to arrive at reliable descriptions of regional and national variability. Several sources of uncertainty in best-estimated constant values, and the state of knowledge in arriving at descriptions of national variability are discussed in detail in Sections 10 and 11of Volume II. Discussed in Section 4.5.4, the aquatic foodweb module uses an internal probabilistic algorithm

US EPA ARCHIVE DOCUMENT

that cycles through the database on prey preferences to randomly select dietary fractions for TL3 and TL4 fish for predicting tissue concentrations (TL - trophic level).

3.4.9 Human Receptor Data

Human receptor points, which include residences and farms, are one of the primary spatial data layers in the 3MRA analysis. They enable human risk to be calculated spatially around a site where people are likely to be located. A geographic information system (GIS) was used to locate these points and collect human receptor numbers and characteristics (e.g., receptor types, age cohorts, etc.) for the 201 study sites. This allows the 3MRA modeling system to develop individual risk distributions around a site that are weighted by population. For the representative national data set, resident human receptor points were located and populated by census block. Farms were located and populated using census block group boundaries, subdivided by farmland use, along with county-level agricultural census data.

Human resident and farmer population data are dimensioned on ring, receptor type, and age cohort. For residents, four receptor types were used to characterize residential exposure and risk: (1) residents; (2); resident home gardeners; (3) resident recreational fishers; and (4) resident home gardener/recreational fishers. For exposure and risk to farmers, the following receptor types were used, depending on whether they are present in the county agricultural census: (1) beef farmers; (2) dairy farmers; (3) beef farmer/recreational fishers; and (4) dairy farmer/recreational fishers. For each of these eight receptor types, five age cohorts were used for the example data set: (1) child1 (infant): younger than 1 year; (2) child2: 1 to 5 years; (3) child3: 6 to 11 years; (4) child4: 12 to 19 years; and (5) adult: 20 years or older. This results in a total of 40 receptor type/age cohort combinations.

Human receptor data were collected on a site-specific basis. Although site-specific data are available for most of the receptor type and land use information necessary to delineate and populate these areas, certain receptor type information (e.g., data on beef and dairy farmers, data on recreational fishers) were only available on the county or state level; these regional data were applied to the site-specific population and land use data to estimate receptor type/age cohort percentages for each subarea.

In general, there are few data gap issues associated with human receptor data; consistent data of an appropriate scale for a national analysis were available for all 201 sites addressed in the analysis. The primary issues and uncertainties are associated with the age of the data and the mismatches based on date and geographic scale. To the extent possible, scale and data discrepancies were addressed using spatial averaging and interpolation to minimize the errors incurred from using different data layers. Larger scale (i.e., block group, county, and state) data were combined with the block data by assuming uniform characteristics across block groups, counties, and states. This step was necessary to allow automated processing of human receptor type and population data. Although this assumption does create some inaccuracies at individual sites, it is assumed valid and appropriate for a national analysis.

US EPA ARCHIVE DOCUMENT

3.4.10 Human Exposure Factors

The human exposure factors used in the 3MRA modeling system are based on national data for these factors provided in the *Exposure Factors Handbook* (EFH) (EPA, 1997d, 1997e, 1997f). These inputs address inhalation and ingestion exposure from contact with media and various food items as well as duration of exposure for the receptor types modeled. Age is an important determinant for most environmental exposure factors, and stratification on age in risk assessment simulations is commonly used to account for this variability. Age groups were selected based on previous precedent and a majority of the data as provided in the EFH, which reduced the need to manipulate the data sets. Site-specific and regional data sets are not available for many of the human exposure inputs, where all human exposure model inputs were collected and processed on a national basis.

The human exposure parameters either are characterized by distributions (stochastic variables) or are fixed values (constants). National distributions were developed for all factors with data that could be used to derive distributions. A few parameters were fixed based on central tendency values from the best available source, either because limited variability was expected or because available data were not adequate to generate national distributions. Section 8.5 of Volume II provides a detailed discussion on various aspects of parameterizing uncertainty distributions, broken-down by source data, distribution type selection, distribution parameter estimation methods, goodness-of-fit tests, parametric versus nonparametric approaches, and specific issues related to the breast milk pathway, which is modeled in 3MRA for Dioxin only.

The human health benchmarks used by the 3MRA are oral reference doses (RfDs), inhalation reference concentrations (RfCs), oral cancer slope factors (CSFs), and inhalation CSFs. The benchmarks are chemical-specific. Human health benchmarks are applied nationally and do not vary by site or region. The human health benchmarks do not vary between receptors (i.e., residents, home gardeners, farmers, and recreational fishers) or age groups. Section 15.5 of Volume II provides a detailed discussion on various, related aspects of uncertainty.

3.4.11 Ecological Receptor Data

The ecological exposure module includes a variety of terrestrial and aquatic habitats that reflect the natural variation of the sites being assessed. The habitats assessed include 14 different terrestrial habitats, including upland, wetland, and waterbody margin habitats, and 8 different aquatic habitats, including cold and warm water streams, lakes, ponds, and wetlands. These habitats cover all types of ecosystems except those associated with estuarine and marine waters, which are not included in the representative national data set. The concept of habitat was chosen as the appropriate level of differentiation for the spatial element of the ecological risk assessments. In this context, the term habitat implies a level of detail and specificity that is meaningful for the exposure scenario at a particular facility site, but does not require extensive biological inventory or field investigation for identification or delineation.

In 3MRA, separation of habitat types into 3 different habitat groups is maintained in the analysis defined by groupings of: (1) terrestrial; (2) aquatic; and (3) wetland. There are 12 habitat types distinguished (5 terrestrial, 6 wetland margin, and 3 waterbody margin). The 5

US EPA ARCHIVE DOCUMENT

terrestrial habitat types are broken down by: grasslands; shrub/scrub; forests; crop fields and pastures; and residential. The 6 wetland margin habitat types are broken down by: intermittently and permanently flooded grasslands; shrub/scrub; and forests, respectively. The 3 waterbody margin habitats are broken down by: rivers/streams; lakes; and ponds. A receptor group is a suite of wildlife species chosen to characterize the exposure scenarios for a particular representative habitat. In 3MRA, separation of data across 9 receptor groups is maintained, defined by: (1) mammals; (2) birds; (3) amphibians; (4) reptiles; (5) soil biota; (6) terrestrial plants; (7) aquatic biota; (8) sediment biota; and (9) aquatic plants.

Finally, ecological receptor populations are further broken-down by 5 trophic level descriptions categorizing various ecological receptors: producers; communities; and trophic levels T1, T2, and T3. In assessing risks to various ecological subpopulations in 3MRA, analysis is constructed along 6 different roll-up levels. Options include roll-up by: (1) distance ring and habitat group; (2) distance ring and habitat type; (3) distance ring and receptor group; (4) distance ring and trophic level; (5) habitat group and receptor group; and (6) habitat group and trophic level. Section 13 of Volume II provides details on the extensive delineation and data collection approaches associated with the various ecological habitats and receptors, imparting additional background information on characterization of uncertainties associated with the data.

3.4.12 Ecological Exposure Factors

Ecological exposure factors are used in to calculate the total exposure dose (in mg/kg/d) to a suite of receptors that consume contaminated media and food items within their respective habitats. The ecological exposure module calculates the exposure dose for 52 terrestrial receptor species as they occur in any of 11 ecological habitats. The body weight and ingestion rate parameters are characterized by fixed values; dietary composition, or the amount of each food item eaten, is characterized by a uniform distribution between reported minimum and maximum values. Prey items were assigned to nine prey categories. The ecological exposure module includes an algorithm to construct a unique, randomly selected diet for each receptor species at each site where it occurs. This algorithm reflects the variability in receptor species' dietary composition. Dietary preference data required by the ecological exposure model includes a list of potential diet items for each species and the maximum and minimum proportion of the species' diet that each item can comprise. Diet items are categorized as one of 17 types of prey.

In general, the ecological exposure factors are for the mean adult body weight for the species; male and female weights are combined, and juvenile body weights are not considered. Dietary composition for each receptor species varies depending on habitat. A database was developed containing all available data relevant to ecological exposures for the 52 selected receptor species. Most of the data were extracted from the *Handbook* (EPA, 1993) and from Sample *et al.* (1997). In some cases, the reported values are the mean of the data collected for a single study; in other cases, the reported values reflect a single measurement. In all cases, all reported values were entered into the database, including means, minima, and maxima. Single reported values were entered as mean values. Data collected from additional sources were treated in the same manner. Imparting aspects of uncertainty, Section 11 of Volume II details various assumptions and approaches used in deriving the exposure factors, broken-down by ingestion rates for: (1) water, (2) food; and (3) surficial soil/sediment.

US EPA ARCHIVE DOCUMENT

For ecological benchmarks, the risk quotient method was used to identify the potential for adverse effects to terrestrial and freshwater receptors. The specific methods used to calculate the protective level (i.e., benchmarks and chemical stressor concentration limits [CSCLs]) varied with the receptor taxa. Protective CSCLs were derived (in ppm) for specific communities and populations in direct contact with contaminated media (i.e., terrestrial plants, soil biota, sediment biota, fish/aquatic invertebrates, herpetofauna). Protective benchmark doses (mg/kg/d) were developed for mammals and birds based on exposure through the food web by ingestion of contaminated food items. The key database developed in support of the ecological risk module was the benchmark/CSCL database, where all the parameters generated are chemical-specific properties. Imparting various aspects of uncertainty, details in how the benchmarks and CSCLs used in the risk quotient were derived are presented in Section 14 of Volume II.

3.4.13 Chemical Properties Database

Chemical properties are required by most of the modules of the 3MRA modeling system. Chemical properties are applied nationally and adjusted based on site-specific or regional pH and temperature conditions. The chemical property values used in the 3MRA were obtained through a combination of modeling, existing databases, and literature review. Certain metal partition coefficients, biodegradation rates, and hydrolysis rate constants were collected through literature surveys for measured values. A summary of chemical properties represented by stochastic distributions is provided in Table 3-2, where all other parameters are defined as constants, adjusted for pH and temperature, where appropriate. While all media temperatures and some pH parameters are generally represented as constant values, these values will change from site to site. As well, pH associated with the aquifer and source terms are described by triangular (probability) distributions. Because many chemical properties are functionally dependent on these parameters, significant variability and uncertainty in chemical properties will be imparted within the national risk assessment. Uncertainty in defining various distribution parameters and constant point-estimates are detailed in references provided in Section 17 of Volume II.

Thermodynamic properties and partition coefficients for organic chemicals were calculated using the SPARC (System Performs Automated Reasoning in Chemistry) model. SPARC calculates a large number of physical and chemical parameters from chemical molecular structure and basic information about the environment (media, temperature, pressure, pH, etc.). Metal sorption isotherms were calculated using the MINTEQA2 geochemical speciation model. MINTEQA2 is an equilibrium speciation model that calculates the equilibrium composition of dilute aqueous solutions in the laboratory or in natural aqueous systems.

Other chemical properties were obtained from literature sources or estimated using empirical methods and expert judgment. Anaerobic and aerobic biodegradation rates and degradation products, for example, were collected using pre-established criteria for the evaluation of field and laboratory studies. The collected biodegradation rates are grouped by pH and temperature regimes. Anaerobic degradation rates are also grouped by redox regime (reducing, sulfate reducing, and methanogenic), which are randomly selected during Monte Carlo runs. Hydrolysis rates were compiled by EPA scientists along with probable pathways and degradation products for the hydrolysis reactions in question. EPA used literature sources where available, and supplemented these published data with laboratory experiments and expertise in

US EPA ARCHIVE DOCUMENT

structure activity relationships as needed. The overall hydrolysis rate constants used by the 3MRA modeling system are calculated as a summation of the rate constants for acid, neutral, and base hydrolysis, using site-specific pH and temperature conditions in the media being modeled.

The metal sorption coefficients for surface soils, surface water, sediments, and wastes were collected from literature where available, or estimated using a combination of empirical relationships with available literature values, geochemical modeling using MINTEQA2, and expert judgment. To represent the nationwide variability in sorption coefficients, the metal Kd values for each media are contained within the 3MRA modeling system as probability distributions that are randomly sampled during model execution.

3.5 Monte Carlo Implementation for the National Assessment

Several key points regarding the science methodology design are revisited next to help provide perspective on the UA/SA plan (Section 9) for evaluating the final 3MRA Version 1.0 technology (i.e., data and modeling system) developed. Taken directly from Marin *et al.* (1999), with minor edits, these points are summarized in the following two sections, and cover objectives and anticipated limitations in conducting a Monte Carlo-based predictive uncertainty analysis. The reader is also referred back to supporting material presented in Section 2.6, and, in particular, materials in Section 2.6.2 taken from Marin *et al.*, (1999) which previously described various potential sources of variability and empirical uncertainty in the 3MRA national assessment.

In conducting the national assessment strategy, a proposed Monte-Carlo procedure to addresses aspects of variability, predictive uncertainty, and sensitivity was designed to meet the following objectives Marin *et al.* (1999):

- *Provide an estimate of the uncertainty in the estimated measures of protection associated with a regulatory waste concentration (C_w);*
- *Provide a mechanism for accounting separately for variability and uncertainty through a two-stage Monte Carlo algorithm;*
- *Provide a (value of information) basis for comparing the potential benefit (reduced prediction uncertainty) versus cost of future sample collection efforts;*
- *Provide a flexible framework that can accommodate alternate policy formulations including different definitions of measure of protection, and both waste and leachate concentration regulatory limits; and*
- *Comply with the U.S. EPA's Guiding Principles for Monte Carlo Analysis (1997a).*

US EPA ARCHIVE DOCUMENT

3.5.1 3MRA Site-Based Approach Monte Carlo (Marin *et al.*, 1999)

The estimation of the measures of protection for receptors of concern for a given waste concentration limit is subject to uncertainty. The proposed approach includes a Monte Carlo algorithm that allows the calculation of uncertainty and variability in the measures of protection. The incorporation of uncertainty in the protection levels allows for the development of regulatory rules that result in more conservative regulatory levels as the uncertainty in risk predictions increases.

This section presents a proposed Monte Carlo structure to support the regulatory framework outlined in Section 3.2, and data collected under ideal conditions outlined in Section 3.4. Although the actual database collected, as outlined in Sections 3.4.1 through 3.4.13, is a departure from ideal conditions, it provides a number of advantages within the data, budgetary and scheduling constraints imposed on the problem. The structure reflects the anticipated compromises made to adjust to limitations associated with the available data and computational constraints, while retaining, to the extent possible, the site-specific and probability sample characteristics of the ideal data set. In particular, the currently available data set consists of a combination of site-specific measurements at existing WMU facilities selected on the basis of a stratified, random sample for selected parameters, together with regional and national databases of surrogate parameters. Specific elements of the 3MRA database include:

1. *A probability subsample of 201 WMU facilities from a stratified sample national survey of WMU facilities (Westat, 1987). This data set provides site-specific measurements for facility characteristics including location and WMU geometries.*

2. *Site-specific evaluations conducted at each of the 201 WMUs in the subsample to determine site-specific parameters.*

3. *Regional databases consisting of non-probability samples of surrogate hydrogeologic parameters and meteorological parameters that allow correlation structures to be established; and*

4. *National databases consisting of non-probability samples of surrogate environmental media characteristics, the (physiologic and behavioral) exposure and response characteristics of the receptors, and the physical, chemical, and biochemical properties of the chemical constituents.*

Given the limitations in the available data, it is anticipated that the initial focus of the Monte Carlo implementation effort will be on significant sampling error sources of uncertainty, and between-site spatial variability of facility/site characteristics. Between-individual variability of receptor characteristics, data measurement errors and model prediction errors will not be addressed initially. They will only be addressed as schedule and resource constraints permit, and as dictated by the results of sensitivity analyses. Additionally, the limitations in the data structure introduce potential non-sampling errors whose magnitude would be difficult to estimate. These errors will not be addressed. As a result, the estimated uncertainties will underestimate the true uncertainties

US EPA ARCHIVE DOCUMENT

The Monte Carlo algorithm will follow a general form of a two-stage Monte Carlo presented. The exact form of the algorithm will depend on the type and amount of available data, the number and types of variability and uncertainties that will be incorporated, and the methods used to model the variability/uncertainty terms. The development of the algorithm will be incremental, moving forward in different stages of refinement as dictated by different testing protocols, including sensitivity analysis and computational benchmarks, and any additional data that may become available in the future.

In addition, the methods used to estimate/model the variability and uncertainty terms will depend on the amount and type of available data, and the computational burden associated with estimation/simulation procedure. It is anticipated that the initial approach will use a combination of empirical and fitted distributions to describe variability [and empirical uncertainty]. Parametric and non-parametric bootstrap methods are available to address uncertainty due to site sampling errors. In all cases, the estimation and reporting of the variability and uncertainty terms will conform to the principles of good practice for the use of Monte Carlo techniques adopted by the U.S. EPA (1997a).

3.5.2 Limitations in Implementation of the Monte Carlo Approach (Marin *et al.*, 1999)

In reality, data limitations, constraints on time and computational resources, and the limits of our scientific knowledge impose a number of departures from the ideal conditions under which to implement 3MRA. First, the physical, chemical, biological and behavioral processes involved are complicated and our knowledge is limited. The required analysis, by necessity, involves a mathematical modeling approximation of the complex causal relationships between waste concentration and the impacts on receptors.

Second, the development of a site-specific model for each facility is impractical. This implied that a generic model had to be developed that could be applied at all sites. A generic model is generally less able to approximate causal relationships than a site-specific model [a generic model does address cause and effect, just at a different level of detail to answer more general questions, or similar questions under conditions of greater uncertainty.] Additionally, computational capacity constraints require that the generic model must be computationally efficient, which forces even greater pressure to make trade-offs between model simplicity and model validity.

Third, resource constraints dictate that the analysis can only be performed at a subset of all of the facilities in the U.S. Ideally, this subset of sites represents a statistically representative sample of the target population, so that inferences from the sample can be extrapolated to all of the facilities in the U.S. However, the sample size will directly affect the uncertainty of the inferred nationwide impacts (see Section 2.6.3).

Fourth, resource constraints also dictate that only a part of the model input data can be collected for all sampled facilities at the site-specific level. The remainder of the model inputs must be characterized through regional and/or national databases, which raises the question of the representativeness of the data to the target population (see Section 2.6.3). Examples of parameters that cannot be practically obtained at the site-specific level for all sites include

US EPA ARCHIVE DOCUMENT

receptor exposure/response physiological and behavioral factors; most hydrogeologic parameters; and climatic characteristics. Finally, computational constraints, data storage requirements, and the spatial resolution of available data impose the need for spatial and temporal averaging at potentially large scales, at all levels of the analysis, including the fate/transport and receptor models.

Under these limitations, additional sources of error will be introduced in the analysis (e.g., errors due to non-representative data), and not all sources of uncertainty or variability (e.g., correlations) can be estimated or identified readily, even in the long run. As a result, estimates of uncertainty in the estimated measures of variability obtained from a two-stage Monte Carlo analysis will only reflect the identified sources of uncertainty for part of the variability. The un-estimated sources of uncertainty will either not be reflected in the uncertainty (e.g., sampling errors, prediction model errors), or remain combined with the variability and not be reflected in the uncertainty (e.g., data measurement errors).

Ultimately, the issue is not whether to incorporate all sources of uncertainty and variability, but rather whether the sources of variability and/or uncertainty that are not included have a significant effect on the regulatory decisions. The key is to include the sources of variability and uncertainty that have the greatest impact on decision-making while meeting the budgetary, scheduling, and computational capacity constraints imposed on the problem.

3.5.3 Summary Points on 3MRA Version 1.0 Implementation

Here we briefly summarize key points regarding implementation of the UA/SA plan for the actual technology developed, as embodied by the 3MRA Version 1.0 modeling system.

- In the original conceptualization of actually developing a screening level technology, it was thought that a relatively fast running, Windows-based modeling system could be constructed to solve the problem statement. The original idea was, in fact, that a technology could be developed that ran on the order of seconds, as opposed to the runtimes of minutes actually associated with the final technology. The latter result was imposed by the requisite science computations needed to satisfy the modeling objectives outlined in Section 3.2 (see also Section 6.6 and Section 9.3.3).

- Runtimes of minutes greatly facilitate use of the modeling system by various users for a variety of inspections, predictive analysis, and in general, model evaluation tasks. Full address of issues of uncertainty and sensitivity analysis for the national scale application across many chemicals and/or metals though, as pointed out in Section 2.5.3, will require large numbers of model runs, necessitating use of multiple PCs to complete tasking in a timely manner.

- The science plan (Marin *et al.*, 1999) covered a wide range of sources of uncertainty and uncertainty analysis techniques with a more modest discussion regarding techniques that might be undertaken for conducting complementary sensitivity analysis, and performance based validation of the technology. In this sense, adequate peer-review of a planned system-level model evaluation approach to address model output sensitivity, predictive

US EPA ARCHIVE DOCUMENT

uncertainty, and performance validation for the national assessment has not been conducted to date. This represents a primary task of the SAB peer-review in progress.

- Two particular aspects of uncertainty generally not presented in detail in the science plan were that of output sampling error and sampling from non-target populations, as in the use of national distributions, for example, that expand the input space beyond that attributable to characteristics of the 201 site (statistically-significant) sample set (Section 2.6.3). It was, of course, critically necessary in this UA/SA plan to fully address these components of total uncertainty in the final analysis, as detailed in Sections 2.5.3, 2.6.6 and 9.

- Statements in Volumes I, II and III that imply the planned 2-stage Monte Carlo has not been fully implemented to date refer only to the lack of treatment of ISE. The "pseudo" 2-stage Monte Carlo to be implemented, as defined in Sections 2.6.3 and 2.6.6 and Section 9, address the capability at this point to separate aspects of variability and uncertainty for the national assessment based on available data sets, and is, in fact, a 2-dimensional analysis. The reader should not confuse this "pseudo" 2^{nd}-stage analysis though as purporting to fully address the original, ideal construct of a two-stage analysis for complete quantification and separation of input sampling error (ISE) in the specification of PDFs, sample measurement error (SME), and model error (ME).

- The approach undertaken here in separating major elements of uncertainty from variability in the existing national assessment, as described in Section 2.6.3 and 2.6.6, embodies a protective strategy in dealing with the limitations of input sampling error associated with the statistical sample of 201 sites (as representing the potential of all existing and future profiles of sets of Subtitle D WMUs). By properly treating the true uncertainty of predicting risks at individual sites in the sample, based on uncertainty imparted through the use of national and regional distributions, the approach consequently captures a great deal of uncertainty that would otherwise be ignored in a completely site-based approach (i.e., direct site-specific measurement of all inputs).

- Two significant aspects not accounted for under the above consideration is the fact that some regional-based groundwater classes (i.e., GWClass 3, 6, 7, 8,and 12) are not represented in the existing site-based database. As well, of the 20 national hydrologic regions, four are not accounted for in the 201site sample (i.e., HydrologicRegion 9, 14, 19, and 20). OSW would need to ensure, through time, that these aspects of regional and national demographics relating to the location of WMUs remain statistically representative. With respect to regional data, the same concern is generally presented in consideration of maintaining existing between-region demographics captured in the sample of 201 sites. This is simply a result of the site-sample design approach that was based on use of existing industrial Subtitle D facilities, which did not inherently account for errors that may arise in capturing future variability in environmental site settings.

- For all data currently represented as national or regional point estimates, predictive uncertainty in model outputs due to random measurement error (RE) is generally not addressed. For point estimates that vary from site to site, it is partially addressed as a

US EPA ARCHIVE DOCUMENT

perspective of hybrid uncertainty convolved in outputs representing "variability" in the pseudo 2nd-order analysis approach outlined in Section 2.6.6.

- Regarding constituent science module error (ME), discussions are presented in Volume I covering: (1) strengths and advantages; and (2) uncertainty and limitations, on a science module-by module basis. 3MRA modeling system error is also addressed in part through: (1) module-level verification, and validation and system-level model comparison activities discussed in Volume III; (2) system-level verification discussed in Section 6 and 9; and (3) performance validation efforts to be undertaken via sensitivity-based performance analysis discussed in Sections 2.8 and 9.

- Due to resource limitations, data needed to employ a full two-dimensional analysis to address ISE of stochastic input distributions could not be feasibly developed to date, which together with model runtime issues associated with the national assessment strategy has, for the time being, precluded full realization of the two-dimensional analysis approach for ISE (Section 2.6.6). Several approaches are summarized in Section 9 that provide perspective on this point, while offering useful abilities to more fully characterize the importance of further separation of SME, and treatment of ISE and ME aspects.

- Extensive address of all SME, ISE, and ME elements of uncertainty in the 3MRA national assessment cannot be feasibly completed in the foreseeable future based largely due to lack of available data and limits of our knowledge. Evaluation of these issues will, of course, remain the subject of ongoing research planning and execution at ORD and OSW. The UA/SA plan proposed in Section 9, however, is intended to provide critical, useful information regarding appropriate prioritization of such efforts.

US EPA ARCHIVE DOCUMENT

Table 3-1. Data Collection Approach by Data Type.

Data Type (Volume II Report Section)	Data Collection Approach		
	Site-Based	Regional	National
WMU (Section 3)	●		●
Waste properties (Section 16)			●
Air Model (Section 16)			●
Meteorological (Section 4)		●	
Watershed and waterbody layout (Section 5)	●		
Surface water (Section 6)		●	●
Soil/vadose zone (Section 7)	●		●
Aquifer (Section 16)		●	●
Farm food chain/terrestrial food web (Section 10)			●
Aquatic food web (Section 11)		●	●
Human exposure factors (Section 8)			●
Ecological exposure factors (Section 12)		●	●
Chemical properties[a] (Section 17)			●
Bio-uptake/bioaccumulation factors[a] (Sections 8, 10, 11)			●
Human health benchmarks[a] (Section 15)			●
Human receptor type and location (Section 9)	●		●
Ecological benchmarks[a] (Section 14)			●
Ecological receptor and habitat type and location (Section 13)	●	●	
Risk and control variables (Section 16)			●

[a]Chemical-specific variables.

US EPA ARCHIVE DOCUMENT

Table 3-2. Chemical Properties Defined by Stochastic Distributions.

Name	Description	Stochastic Distribution Type
ChemAerBioRate	Aerobic Biodegradation rate	Triangular (Hg); Uniform organics where applicable; no data other metals)
ChemAnaBioRate	Anaerobic Biodegradation rate	Triangular (Hg); Constant or Uniform organics where applicable; no data other metals)
ChemAnaRedRate	Anaerobic Reduction	Triangular (Hg only)
ChemKd	Partition Coefficient for Media	Triangular
ChemKDoc	Koc	Triangular (Hg); Constant others
ChemMetBioRate	Anaerobic Biodegradation under Methanogenic Red.	Uniform, Constant, or Empirical (organics only)
ChemPh	pH assumed for these properties	Triangular (same as SrcPh)
ChemSO4BioRate	Anaerobic Biodegradation rate under SO4 Reduction	Uniform, Constant, or Empirical (organics only)
ChemTemp	Temperature assumed for these properties	Constant (by site; same as SrcTemp)
ChemWDiff	Water Diffusion Coefficient	Uniform (except Hg; Constant)

US EPA ARCHIVE DOCUMENT

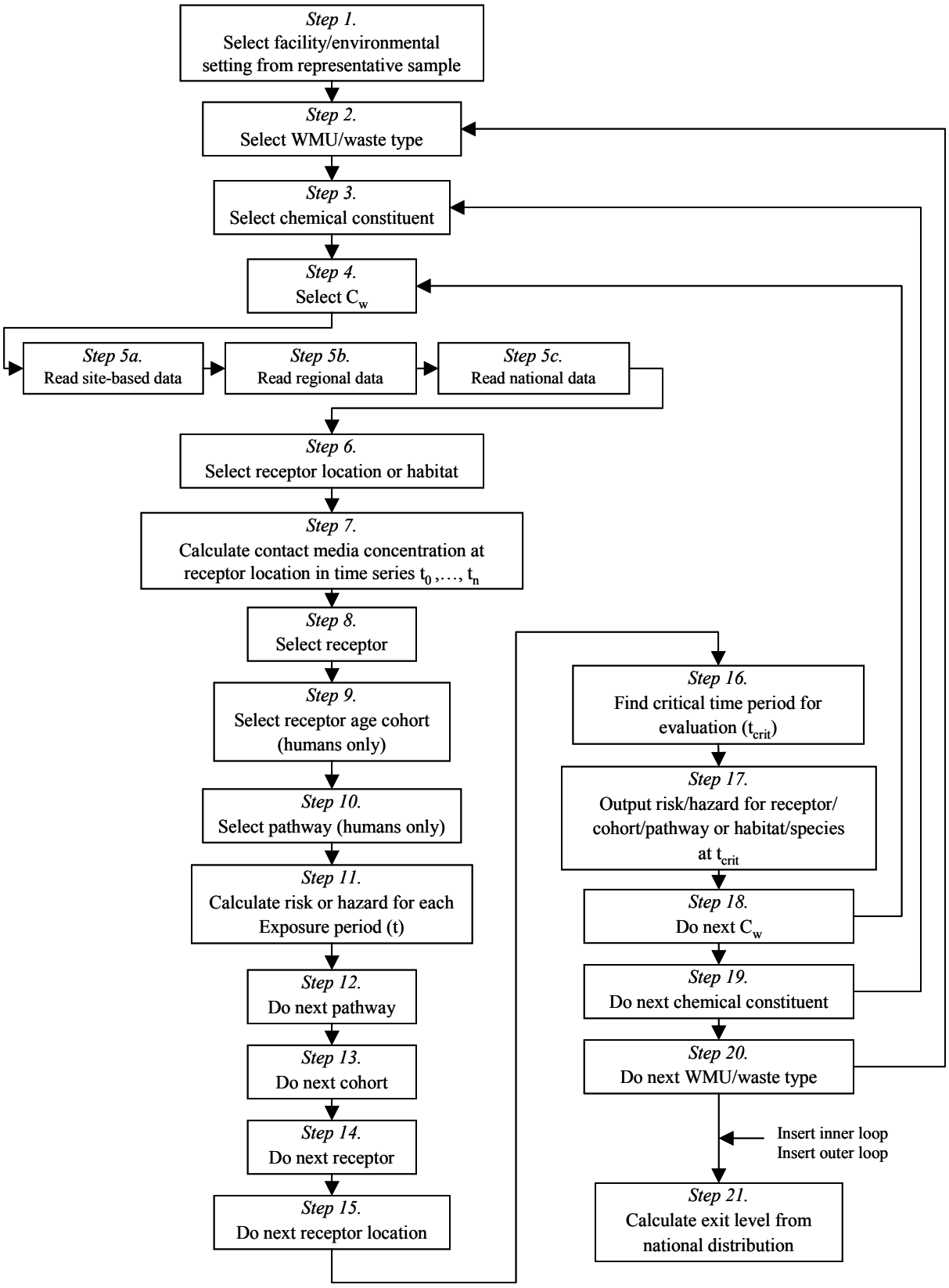

Figure 3-1. Conceptual Implementation of 3MRA Model.

US EPA ARCHIVE DOCUMENT