A FRAMEWORK FOR FINITE-SOURCE MULTIMEDIA, MULTIPATHWAY, AND MULTIRECEPTOR RISK ASSESSMENT 3MRA

Work Assignment Manager
and Technical Direction:

Dr. Zubair A. Saleem
U.S. Environmental Protection Agency
Office of Solid Waste
Washington, DC 20460

Prepared by:

HydroGeoLogic, Inc.
1155 Herndon Parkway, Suite 900
Herndon, VA 20170
Under Contract No. 68-W7-0035

U.S. Environmental Protection Agency
Office of Solid Waste
Washington, DC 20460

July 1999

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

A number of individuals have been involved with this development. Dr. Zubair A. Saleem of the U.S. EPA, Office of Solid Waste, provided overall technical direction and review throughout this work. This report was prepared primarily by Dr. Carlos Marin of Ambiotec, Inc., and Dr. Varut Guvanasen of HydroGeoLogic, Inc. Useful inputs provided by many individuals working on the Hazardous Waste Identification Rule (HWIR) are greatly appreciated. They include: Messrs. Barnes Johnson, Stephen Kroner, Gerry Laniak, David Cozzie, and Drs. David Brown and Sharon LeDuc, among many others. We also appreciate the comments on the assessment strategy by the independent peer reviewers: Drs. Mitchell Small, Yoram Cohen, and Paul Deisler.

DISCLAIMER

The work presented in this document has been funded by the United States Environmental Protection Agency. Mention of trade names or commercial products does not constitute endorsement or recommendation for use by the Agency.

1.0 INTRODUCTION

The U.S. EPA developed a methodology to set regulatory threshold levels for chemical constituents in wastes based on the expected groundwater impact of constituents leaching from Subtitle D waste management units (WMUs). On March 29, 1990, the Agency applied the methodology to develop the final toxicity characteristic (TC) levels (55 FR 11798, March 29, 1990). These levels are used to identify those wastes, defined as hazardous, which are subject to the Resource Conservation and Recovery Act (RCRA) Subtitle C regulations. They are defined as hazardous due to their potential to leach significant concentrations of specific toxic constituents. In this approach, if a waste with constituents having concentrations, determined from the Toxicity Characteristic Leaching Procedure (TCLP), exceed any of the corresponding TC regulatory limits; the waste is classified as toxic hazardous waste. The Agency assumed that the waste in the landfill behaves as an infinite source of a contaminant in the leachate in the TC modeling approach. The Agency revised and enhanced the TC modeling approach to accommodate finite-source conditions, as well as contaminant transformation and sorption processes. The revised and enhanced modeling procedure has been incorporated into the EPA's Composite Model for Leachate Migration with Transformation Products model (EPACMTP; EPA, 1996).

In December of 1995 the Agency proposed to amend existing regulations for disposal of listed hazardous wastes under RCRA. The 1995 proposal (60 FR 6634, December 21, 1995) outlined the Hazardous Waste Identification Rule (HWIR) that was designed to establish constituent-specific exemption levels for low risk solid wastes. Wastes applicable 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 exemption level criteria defined by the HWIR methodology, would no longer be subject to the hazardous waste management system specified under Subtitle C of RCRA for those wastes. Basically, this established a risk-based "floor" for low risk hazardous wastes that would encourage waste minimization, and the development of innovative waste treatment technologies. The purpose of the rulemaking was to reduce possible over-regulation arising from the older "mixture" and "derived-from" rules promulgated earlier. Note that, in a number of cases, wastes were listed on the basis of containing both toxic hazardous constituents and exhibiting one or more of the hazardous waste characteristics that do not relate to chemical toxicity (e.g., ignitability, corrosivity, reactivity). If such a waste still exhibits any characteristic after complying with the exemption criteria proposed in the HWIR, it must continue to be managed as a characteristically hazardous waste.

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 they apply regardless of the concentration or mobility of 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.

Extensive reviews of the original proposal were conducted by the EPA Science Advisory Board (SAB), the Office of Research and Development (ORD), and numerous industrial and environmental stakeholders. The collective conclusion resulting from the reviews was that the technical basis of the proposed 1995 rule was not sufficient to allow the rule to be promulgated.

As part of a consent decree entered in the U.S. District Court on April 7, 1997, the EPA agreed to an accelerated schedule for the development of a methodology and for the promulgation of a revised HWIR. The decree specified that the Agency will propose a rule by October 31, 1999 and will finalize the rule by April 30, 2001 [ETC vs Browner, CA No. 94-2119 (DDC 1997)].

In the HWIR95 framework (U.S. EPA, 1995), groundwater and non-groundwater pathways were separately analyzed. A review of this approach by the EPA's Science Advisory Board indicated that 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 cannot 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 (SAB, 1995). In response to the SAB's recommendations, a consistent multimedia, multireceptor, multipathway risk assessment (3MRA) approach was conceptually formulated. A conceptual risk assessment procedure for determining waste concentration limits for hazardous chemicals using a finite-source 3MRA approach is outlined in this report. The approach includes a Monte Carlo algorithm that allows ultimately the calculation of the uncertainty in the regulatory levels as a function of input parameter sampling and measurement errors, and prediction model errors.

The assessment strategy may be implemented fully provided that all of the required components and necessary resources are available. However, the scope of its actual implementation will depend on the availability of data, computational resources, and time constraints.

2.0 CONCEPTUAL FRAMEWORK FOR 3MRA

The objective of this section is to outline a conceptual framework for the finite source, multimedia, multipathway, multireceptor risk assessment (3MRA) for HWIR99. Details of the implementation, including fate/transport modeling algorithms, data collection and analysis efforts, and computational efficiency considerations are currently under development and are not addressed here.

The outline of the proposed approach is not meant to be definitive, but is rather intended to provide a foundation for the development of the modeling framework. There are a number of options available for the regulatory objectives including, for example, issues associated with the definition of protection measures, definition of the problem statement, and the number and type of receptors evaluated, to name just a few. While the details of the framework will depend on the specific assumptions adopted (see Section 3.0), the proposed framework is intended to be sufficiently general to accommodate alternative options.

2.1 OBJECTIVES OF 3MRA

The objective of the multimedia, multipathway and multireceptor risk assessment (3MRA) framework is to develop regulatory waste concentration limits for chemicals in wastes managed in industrial Subtitle D waste management units. The concentration limits apply uniformly throughout the U.S. and are chemical specific. A waste with measured concentrations of all regulated chemicals that are lower than their regulatory limits is defined as non-hazardous and can be managed in a Subtitle D unit. A waste with concentrations of any chemical constituent higher than any of the regulatory limits is defined as hazardous and must be managed in a more stringently regulated Subtitle C facility.

The regulatory waste concentration limits are determined by evaluating the impact that wastes managed in Subtitle D units could have on human and ecological "receptors of concern". The impact to human and ecological receptors is quantified in terms of "measures of protection" based on risk and is evaluated at a nationwide level.

The primary measure of protection used in this document is the percentage of the nationwide receptors of concern that are exposed to cancer risks that exceed specified risk (cancer risk/hazard quotient) target levels. There are a number of alternative measures of protection that can be used in the assessment framework. Alternative measures include variations of the primary protection measure based on the type of receptor(s) protected, as well as measures that can be used either independently or in conjunction with the primary protection measure. Some of these alternatives include, for example, secondary protection measures based on the percentage of protected receptors for subpopulations; the expected number of excess lifetime risks among the nationwide population of receptors of concern; or the average population weighted risk of receptors exposed to risk levels above given limits. The assessment framework has been designed around the primary measure of protection. However, it retains the flexibility to address different measures of protection by providing an output data base that can be queried off-line in alternative ways depending on the adopted measures of protection.

The fundamental question that the proposed framework is designed to answer for a given chemical can be stated in the following way: *if a "receptor of concern" is defined as all receptors of a given type that currently reside within a specified radius of all currently existing Subtitle D waste management units in the continental U.S., then what percent of the total number of current receptors of concern would be exposed to risk/hazard quotient levels above specified target levels if each facility were to manage the chemical at the same concentration at all facilities*?

Clearly, any attempt to determine nationwide risks is a challenge. Performing a risk assessment at a site-specific level is difficult enough. Extending the site-specific effort to a nationwide scale introduces an additional, and significant, layer of difficulty. Ideally the nationwide risk assessment would be performed in four steps:

1. Identify all current Subtitle D waste management units in the continental U.S.;
2. Collect all of the site-specific data necessary to characterize each unit and associated site/receptor characteristics, and relevant processes;
3. Develop a site-specific mathematical model to predict the impacts at each site; and
4. Run the site-specific model at each of the sites to predict the nationwide impacts to the "receptors of concern".

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. *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 to receptors. *Second*, the development of a site-specific model for each unit is impractical. This implies that a generic model will need to be developed that can be applied at all sites. A generic model is generally less able to approximate causal relationships than a site-specific model. Additionally, scheduling 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 the 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. *Fourth*, resource constraints 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 data bases, which raises the question of the representativeness of the data to the target population. Examples of parameters that cannot be practically obtained at the site specific level for all sites include receptor exposure/response physiologic and behavioral factors, most hydrogeologic parameters, and climatic characteristics. Finally, computational constraints 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.

These are some of the principal departures from the ideal conditions that will be required for the analysis. Each departure forces a tradeoff between the uncertainty of the estimated impacts and the need to incorporate simplifications in the analysis. Where to simplify and how to simplify are difficult questions to answer. Ultimately decisions on simplification will be based on whether the tradeoffs affect the answer in a significant way. This points to the need to incorporate flexibility in the approach to determine the sensitivity of the inferences to simplifying assumptions; and the need to quantify uncertainty in the inferences. It also points to the need to develop the methodology first without regard to the limitations imposed by the various constraints so that any departures from the ideal conditions can be measured and guided from this starting point.

The framework incorporates a Monte Carlo algorithm that allows the calculation of the uncertainty and associated confidence levels in the estimated impact measures as a function of the sampling and measurement errors of input parameters, and the errors in the prediction models. The modeling procedure uses a forward calculation and maintains mass balance at the source. It can be used to estimate the relative importance of controllable sources of uncertainty; and can also be used to develop regulatory schemes that can result in more conservative limits as the uncertainty in the estimated measures of protection increases.

2.2 PROBLEM STATEMENT

The underlying premise of the HWIR is that there are wastestreams currently included in hazardous waste listings that are effectively non-hazardous, that is, the wastestreams, if disposed of in accordance with Subtitle D regulations (as opposed to the hazardous waste requirements of Subtitle C), would not pose a significant health threat to human and ecological receptors. To quantify specific criteria for determining which wastestreams may "safely" exit the hazardous waste disposal program, the Agency must perform a technical assessment of the potential health risks related to the reduced requirements of Subtitle D disposal. The primary criterion for HWIR is related to the concentrations of HWIR chemicals-of-concern exemption in wastestreams. Wastestreams containing chemicals with concentrations below Agency specified thresholds would exit the hazardous waste system. Conversely, those wastestreams containing concentrations of any HWIR constituent above the chemical-specific threshold would remain in the hazardous waste program.

Given this background, the HWIR technical assessment problem can be defined as:

> *Problem Statement: To determine constituent-specific wastestream concentrations that represent a threshold below which Subtitle C disposal will not be required and thus the wastestream may "exit" the hazardous waste management system and can be managed in a Subtitle D (non-hazardous) waste management system.*

The HWIR is a risk-based rule, thus the constituent-specific waste exemption levels are 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. Also, the HWIR is a national ruling, thus the exit levels must apply to all wastestreams under all Subtitle D waste management scenarios.

Thus, the HWIR99 assessment will be a screening-level risk-based assessment of potential human and ecological health risks resulting from long-term (chronic) exposure to HWIR chemicals released from land-based waste management units (WMUs) containing currently 'listed' waste streams. The assessment of potential health risks will be conducted for both human and ecological receptors. The assessment will be national in scale and site-based, that is, risks will be assessed at individual sites across the U.S. where HWIR WMU's may be located. 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 from the WMU to each environmental medium, the fate and transport of the chemical through a multimedia environment, and the receptor-specific exposures and risks will be simulated. 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 includes an estimation of the potential exposures per exposure pathway/receptor and aggregated across pathways followed by an estimate of the resulting carcinogenic and noncarcinogenic health effects. The end point of the technical assessment is a compilation of the risks to form a national scale joint distribution reflecting the relationship between chemical concentration in wastestreams 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.

3.0 RISK ASSESSMENT METHODOLOGY: TECHNICAL APPROACH

The conceptual foundation of the technical approach to achieving the HWIR99 goals is the risk paradigm and the associated relationship between a source of contaminant, its release to and transport through the environment, subsequent contact (i.e., exposure) with human and ecological receptors, and the resulting risk of health effects.

The following subsections are organized to describe the essential features of the proposed site-based risk assessment methodology, 3MRA.

3.1 TECHNICAL BASIS

To provide a risk assessment context to the discussions that follow, Table 3.1 presents the dimensions of the proposed HWIR99 Integrated Multimedia Risk Assessment, and Figures 3.1 and 3.2 depict the exposure pathways for human and ecological receptors, respectively.

3.1.1 Assumptions

There are a number of key decisions and assumptions regarding the assessment of exposures and risks as well as the methods for establishing national exemption criteria and for driving the national threshold levels. These decisions and assumptions reflect the manner in which certain requirements of the assessment will be satisfied, e.g., an Agency policy decision or a technical assumption based on scientific judgment. They are presented to help frame the presentation of the detailed technical approach in subsequent sections.

- Regulatory threshold waste concentration limits are determined for each constituent of concern. The threshold levels are based on the evaluation of the nationwide impacts to the health of receptors of concern resulting from the management of the given chemical in Subtitle D units located throughout the U.S.

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

- A waste concentration limit for each WMU type is derived independently for each chemical.

- Nationwide impacts are determined by aggregating the impacts of individual waste management facilities. The individual waste management facilities represent actual sites located throughout the U.S. The location and physical characteristics of the sites (e.g., surface area, volume, number and type of waste management units) were determined from a statistical sample of industrial Subtitle D facilities. The sample represents a "snapshot" of industrial waste management facilities for the year 1986 throughout the U.S. The resulting impacts for the individual facilities in the sample are aggregated to be national level. This is done by extrapolating the sample of sites to the population of sites in the U.S. by using the sampling weights associated with each individual facility.

Table 3.1
Multimedia Risk Assessment Dimensions

CONTAMINANTS
- Organics (approx. 200)
- Metals (20)

SOURCE TYPES
- Landfill
- Land Application Unit
- Surface Impoundment
- Aerated Tank
- Waste Pile

SOURCE TERM CHARACTERISTICS
- Mass Balance
- Multimedia Partitioning
- Chemical Decay

SOURCE RELEASE MECHANISMS
- Erosion
- Volatilization
- Runoff
- Leaching
- Particle Resuspension

TRANSPORT MEDIA
- Atmosphere
- Soil
- Vadose zone
- Saturated zone
- Surface water

FATE PROCESSES
- Chemical/Biological Transformation (and associated products of transformation)
- Linear partitioning (water/air, water/soil, air/plant, water/biota)
- Nonlinear partitioning (metals in vadose zone)
- Chemical Reaction/Speciation

AGE GROUPS FOR HUMAN RECEPTORS
- Infant < 1 year
- Child-a 1- 5 years
- Child-b 6 - 11 years
- Child-c 12- 19 years
- Adult 20+ years

INTERMEDIA CONTAMINANT FLUXES
- Source -> Air (vol, resuspension)
- Source -> Vadose zone (leaching)
- Source Surface soil -> Local Watershed Soil (erosion, runoff)
- Air -> Watershed/Farm /Habitat Soil (wet/dry dep)
- Air -> Surface water (wet/dry dep)
- Air -> Vegetation (dep/uptake)
- Farm/Habitat Soil -> Vegetation (root uptake)
- Watershed Soil -> Surface water (erosion, runoff)
- Surface water -> Aquatic organisms (uptake)
- Surface water -> Sediment (sedimentation)
- Vadose zone -> Groundwater (percolation)
- Groundwater -> Surface water
- Soil -> Vegetation (uptake, dep)
- Vegetation, Soil, Water -> Beef and dairy (uptake)

.....

FOODCHAIN
- Human (Farm)
- Human (Aquatic)
- Ecological (Aquatic Habitat)
- Ecological (Terrestrial Habitat)

RECEPTORS

Human
- Resident (Adult & Child)
- Beef Farmer (Adult & Child)
- Dairy Farmer (Adult & Child)
- Home Gardener (Adult & Child)
- Recreational Fisher (Adult & Child)

Ecological

Mammals, Birds, Soil Communities, Terrestrial Plants, Aquatic Communities, Benthic Communities, Aquatic Plants, Amphibians, Herpes, and Reptiles.

EXPOSURE ROUTES
- Ingestion (plant, meat, milk, aquatic food, water, soil)
- Inhalation (gases, particulates)
- Direct Contact (soil, water)

HUMAN AND ECOLOGICAL RISK ENDPOINTS
- Human Cancer Risk
- Human Noncancer Hazard Quotient
- Ecological Population and Community Hazard Quotients

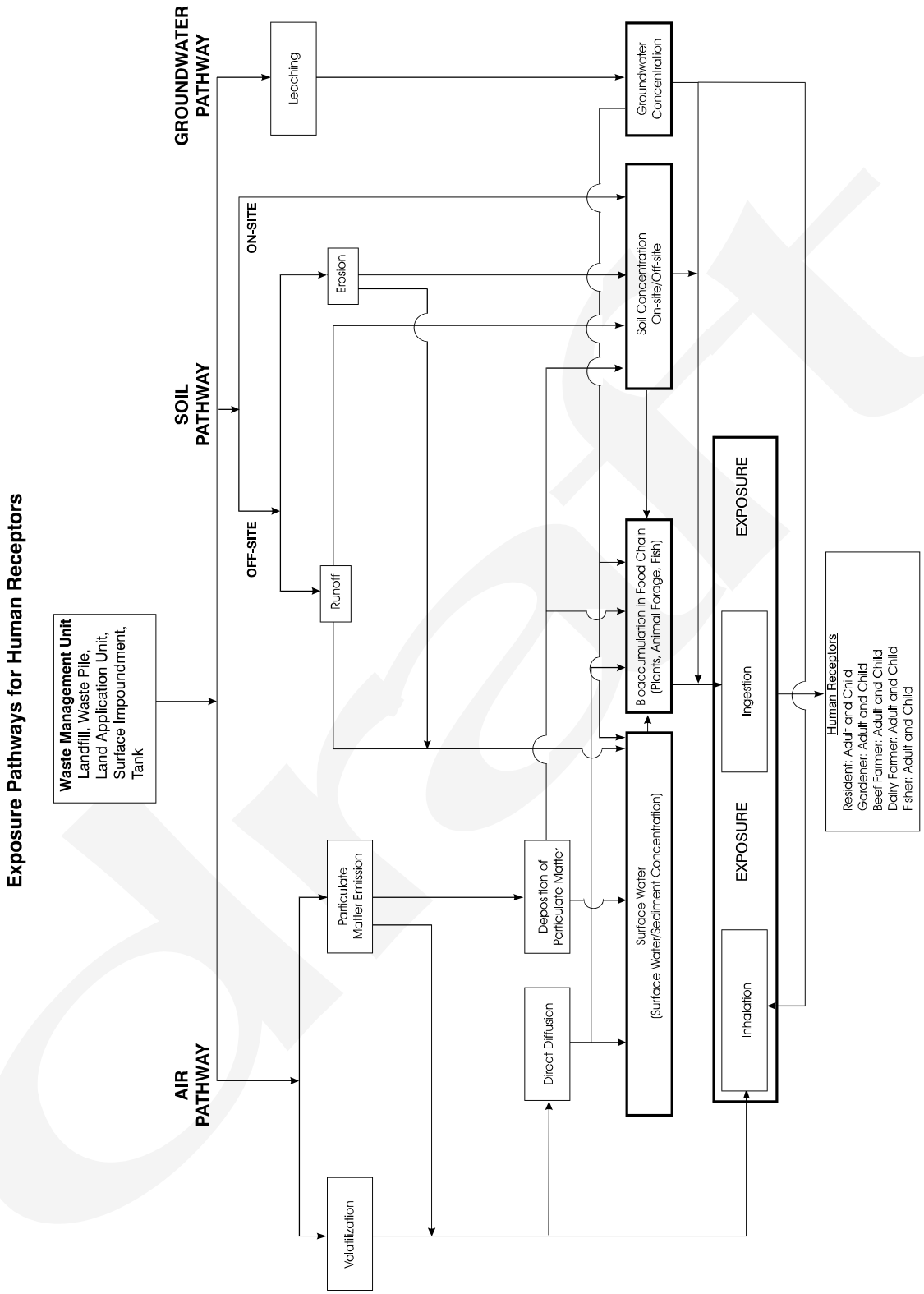

Figure 3.1 Exposure pathways for human receptors.

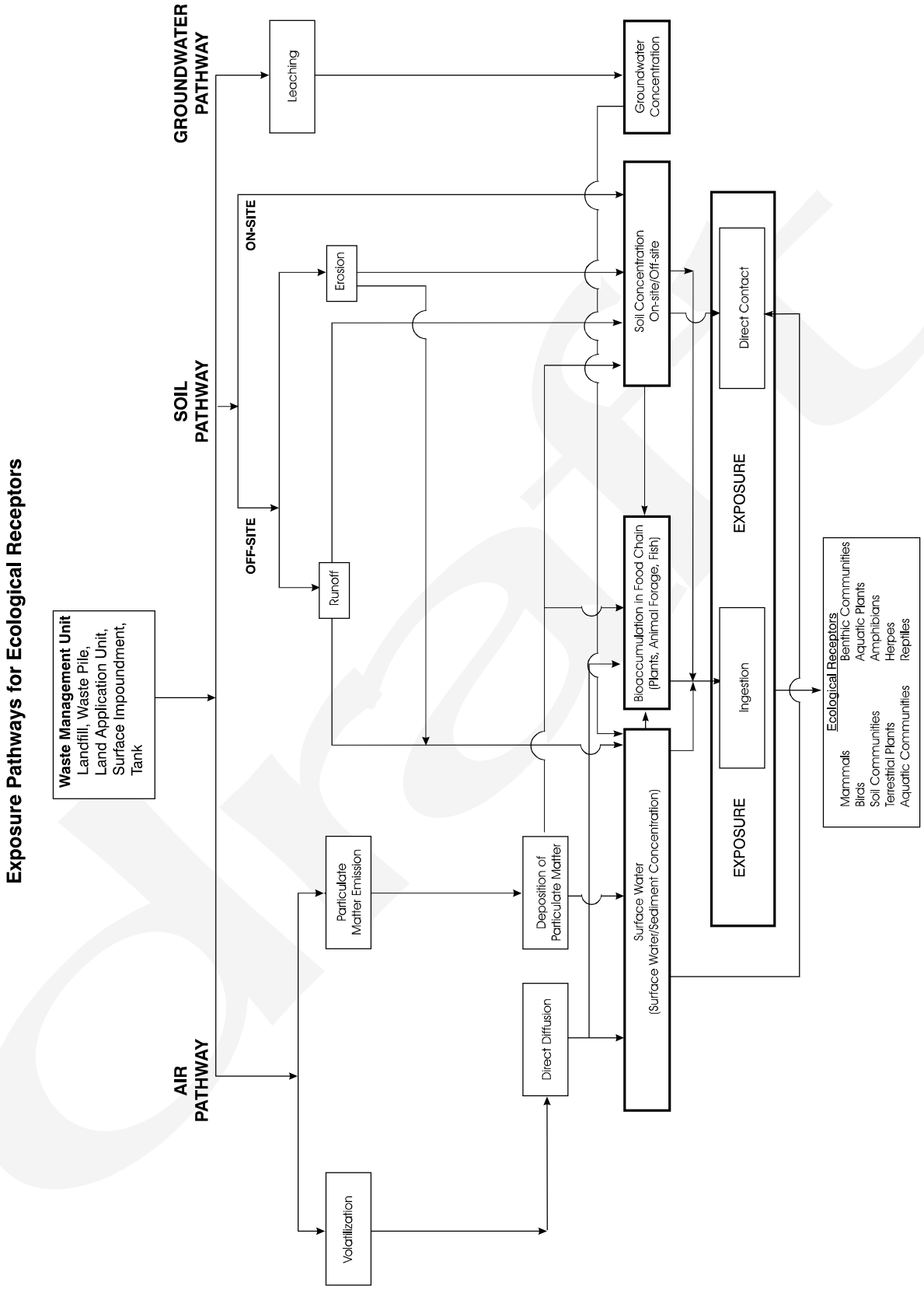

Figure 3.2 Exposure pathways for ecological receptors.

- 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.

- The approach is designed to 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 will be used.

- A waste management facility (WMF) can contain more than one type of waste management unit (WMU), and more than one WMU of each type. However, the impact of the sources is represented by modeling a single unit with source characteristics (e.g. area and volume) given by the average of the individual units. This unit is located at the centroid of the WMUs of the same type.

- The assessment of the impact of a single waste management facility on receptors of concern is based on the consideration of near-field, long-term (chronic) impacts from the operation and closure of the waste management unit.

3.1.2 Analytic Bases

In addition to the assumptions listed above, the risk assessment is based on the following:

- The effects of the different WMU types within a site are considered separately in the exit-level determination decision context.

- A given site is defined by the area contained within a two-kilometer distance from the unit boundary as defined by the unit area. Receptors and other major features of the site (e.g., rivers, lakes) are located by data described in the section on Assessment Data.

- 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.

- Receptors of concern include both human and ecological receptors that are located within a given distance (e.g., two km) of the waste management unit. Both human and ecological receptors include receptors of different types (see Tables 3.1 and 3.2).

- 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 is not being considered because of limited data as compared with those available for ingestion and inhalation and has been excluded from the risk assessment for HWIR99. Exposure media for human receptors include groundwater, soil, air, biota (vegetables, meat, dairy products, etc.).

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

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

- The total mass of a given chemical constituent that is managed in a WMU is a finite value, M_{Total}. The value of M_{Total} can be different for different unit types in a facility, but the waste concentration, C_w, is the same for all unit types in all facilities The total volume V_{Total} in which M_{Total} resides must not exceed the total available capacity of the WMU which is a measured site-specific quantity. The total mass of a constituent can be accumulated incrementally throughout the WMU's operating life.

$$M_{Total} = M(t_0) + \int_{t_0}^{t_{oplife}} Q_m(t)dt \qquad \textbf{(1)}$$

where: $M(t_0)$ = Mass of given chemical at time t_0
M_{Total} = Total mass of given chemical
t = Time ϵ $[t_0, T_{max}]$
t_{oplife} = Unit operating life
C_W = Incoming waste concentration of given chemical
C_{Wout} = Concentration of given chemical in the waste volume removed from the WMU
ρ_{HW} = Density of regulated (hazardous) waste.
ρ_{Wout} = Density of the waste volume removed from the WMU.
$Q_m(t)$ = Net rate with which waste mass of given chemical is changed in the unit
= 0, when $t > t_{oplife}$
$$= \left(C_w \, \rho_{HW} \, \frac{d}{dt} V_{Win} - C_{Wout} \, \rho_{Wout} \, \frac{d}{dt} V_{Wout} \right), \; when \; t \le t_{oplife}$$

V_{Win} = Waste volume entering the WMU at time t
V_{Wout} = Waste volume removed from the WMU at time t

For some types of WMU, such as landfill, the total available storage capacity of the WMU is assumed fixed, and no additional mass beyond the storage capacity (initial finite) value can be added to the WMU (see Equation (2) below). Whereas for other types of WMUs, e.g., waste piles, waste mass is removed and replenished periodically during the operational lifetime of the WMU.

$$V_{Total} = V(t_0) + \int_{t_0}^{t_{oplife}} Q_V(t)dt \le S_{WMU} \qquad \textbf{(2)}$$

where: $V(t_0)$ = Waste volume at time t_0
V_{Total} = Total unit volume
$Q_V(t)$ = Net rate with which waste volume is changed in the unit

$$= \frac{d}{dt}\left(V_{Win} - V_{Wout}\right)$$

= 0, when $t > t_{oplife}$
S_{WMU} = WMU's available storage capacity

- Mass balance in the waste management unit is 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.

  Mass balance in the proposed approach is based on the following mass conservation equation:

$$\frac{dM(t)}{dt} = Q_m(t) - \sum_{ipr=1}^{N_{Pr}} R_{ipr}(t) \qquad \textbf{(3)}$$

  Integrating the above equation yields,

$$M(t+\Delta t) = M(t) + Q_m(t)\Delta t - \sum_{ipr=1}^{N_{pr}} R_{ipr}(t)\Delta t \qquad \textbf{(4)}$$

  where $R_{ipr}(t)$ = Rate with which mass is released through process ipr
  Δt = Time step size; and
  N_{pr} = Number of physico-bio-chemical processes by which the mass is released from the waste management unit (see source release mechanisms, Table 3.1).

- Some of terms used in this document are defined below for ready reference:

  Exposure pathway - The course a chemical takes from the source(s) to an exposed organism. Each exposure pathway includes a source, an exposure route, a contact medium, and the location of a representative receptor in an exposure area.

  Contact medium - The substance that transports the constituent(s) from the source to an exposed organism. Contact media include, for example, surface water, groundwater, air, and soil.

  Exposure route - The manner in which a chemical(s) come(s) into contact with, or introduced into, an exposed organism. For example, exposure routes include inhalation, and ingestion.

  Exposure area - The area in which receptors are located. In each exposure area, receptors of the same type are replaced by a representative receptor randomly located anywhere in the exposure area.

- 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}. The contaminant concentration for any contact medium at any point within an exposure area at a given time is given by the areal average over the exposure area.

- Each receptor type in an exposure area at a site is 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 each successive cohort are lagged by one-year interval from time t_0 to T_{max}. If t represents time in years, we can refer to each longitudinal cohort uniquely as cohort t. Each cohort t is assumed to be exposed to annual contact medium concentrations from the age of a years to $a + d_{fgh}$ years (from time t to time $t + d_{fgh}$), where d_{fgh} is the total exposure duration defined in Equation (5) below. Each cohort is allowed to age naturally and is immediately preceded and followed by two identical cohorts t - 1 and t +1, respectively, with $t = t_0 , t_0 + 1, \ldots, T_{max}$. Each receptor type for humans represents a distinct age group (the age at which exposure begins) and for each age group there is a series of cohorts. With the exception of the exposure concentration, the characteristics (e.g., exposure characteristics; location of receptors; exposure areas; number of receptors in an exposure area) of each cohort are the same. In general, each receptor type is assumed to reside within the exposure area during the exposure duration. A nationwide probability distribution based on the EPA's Exposure Factors Handbook data (U.S. EPA, 1997) will be used to simulate the exposure factors for each receptor type. The exposure factors for ecological receptors will be based on the EPA's Wildlife Exposure Factors Handbook (U.S. EPA, 1993).

- The response variables (e.g.- cancer slope factor, (CSF) of the dose which is a population-level benchmark) can vary between chemicals, receptor types (ecological receptors only), and exposure routes, but are assumed as a matter of policy to not exhibit variability (or uncertainty) between individual receptors of the same type within an exposure area.

- The impact to each receptor type is evaluated in terms of risk measures (e.g., risk or hazard quotient) that provide a measure of the impact. In the case of carcinogens, (or non-carcinogens where inhalation and ingestion act on the same organ) the individual exposure route risks can be aggregated to estimate the aggregate risks.

- Risks can be described by pathway, media, and exposure route. If there are:

  b = 1, 2…, nb(f) WMU types;
  e = 1, 2…, ne chemicals;
  f = 1, 2…, nf sites;
  g = 1, 2…, ng(f) exposure areas;
  h = 1, 2…, nh(f) receptor types;
  i = 1, 2,…, ni(k) contact media;
  j = 1, 2,…, nj(i,k) pathways associated with exposure route k and contact medium i;
  k = 1, 2,…, nk exposure routes, and

  for a representative receptor of type h in exposure area g of site f, then the pathway, media and exposure route risks are defined as follows:

1) The pathway specific risk ($PR_{befghijkt}(C_w)$) for chemical e at waste concentration $C_{w,}$ for an individual cohort t (that starts exposure at time t) associated with representative receptor type h for pathway j, involving exposure route k and contact medium i, in exposure area g, in WMU of type b in site f is given by the sum of the concurrent doses (doses in the same exposure period) to the receptor during exposure duration d_{fgh} :

$$PR_{befghijkt}(C_w) = \sum_{T=t}^{t+d_{fgh}} \frac{C_{befgijkT} \cdot I_{fghikT} \cdot EF_{fghikT} \cdot \beta_{ehk} \cdot \delta_d}{\Delta_e \cdot W_{fghT} \cdot 365} \quad \textbf{(5)}$$

where:

$C_{befgijkT}$ = Annual concentration of constituent e, in contact medium over the exposure area associated with exposure route k and pathway j in exposure area g of site f in year T due to waste concentration C_w in WMU type b.

I_{fghikT} = Daily intake (kg/day) of contact medium i associated with exposure route k and pathway j by cohort t associated with a representative receptor of receptor type h in exposure area g of site f in year T

Ef_{fghikT} = Exposure frequency (days/yr) for cohort t associated with a representative receptor of receptor type h from media i associated with exposure route k in exposure area g of site f in year T

d_{fgh} = Exposure duration (yrs) for cohort t associated with a representative receptor of receptor type h in exposure area g of site f in year T

β_{ehk} = Carcinogenic risk potency (and the inverse of the reference dose RfD for non-carcinogens) for exposure route k for chemical e for cohort t associated with individual receptor m of receptor type h $(mg/kg/day)^{-1}$

Δ_e = Averaging time for chemical e (yrs)

W_{fghT} = Body weight for cohort t associated with a representative receptor of receptor type h in exposure area g of site f in year T

δ_d = Time step (1 year).

Figure 3.3 shows the relationship between the annual concentration in contact medium i associated with exposure route k and pathway j, $C_{befgijkt}$ (obtained from fate and transport component modules), and the pathway risk associated with the contact medium for a single cohort, $PR_{befghijkt}$. The figure demonstrates that the pathway-specific risk for a single longitudinal cohort t_1 is based on medium concentration averaged between time t_1 and time $t_1 + d_{fgh}$.

2) The contact medium specific risk ($MR_{befghikt}(C_w)$) for media i associated with exposure route k from chemical e for cohort t associated with representative receptor type h in exposure area g of site f for waste concentration C_w in WMU type b is given by the sum of the concurrent individual pathway risks corresponding to the exposure route and contact medium:

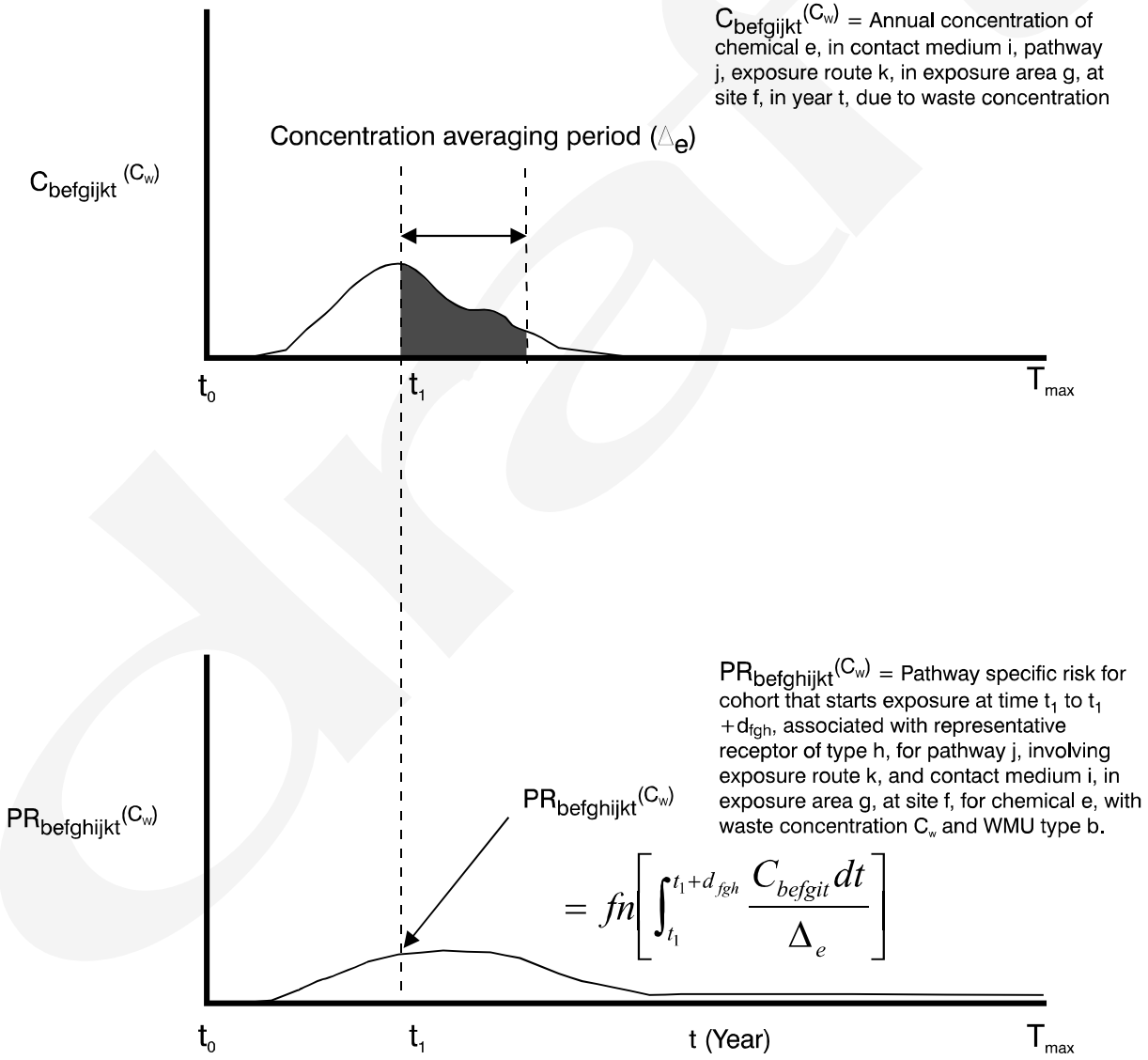

Figure 3.3 Relationship between exposure concentration and pathway risk.

$$MR_{befghikt}(C_w) = \sum_{j=1}^{nj(i,k)} PR_{befghijkt}(C_w) \quad \textbf{(6)}$$

The pathway risk, as described in Equation (5), is presented here in the event that the determination of pathway risks is necessary or of interest. However, for HWIR99, pathway risks may not be calculated due to the anticipated difficulties resulting from the computer storage and computational constraints.

3) The exposure route specific risk ($ER_{befghkt}(C_w)$) for a receptor cohort t associated with representative receptor of type h for exposure route k at time t, in exposure area g, for chemical e at site f for waste concentration C_w in WMU type b is given by the sum of the concurrent risks of the cohort from each media i associated with exposure route k over the exposure duration d_{fgh}:

$$ER_{befghkt}(C_w) = \sum_{i=1}^{ni(k)} MR_{befghikt}(C_w) \quad \textbf{(7)}$$

- There are two types of aggregate risks, for cohort t of receptor type h, which are of interest: aggregate contact medium-specific risk, and aggregate receptor-specific risk.

  In the case of a carcinogen (or a non-carcinogen where the exposure routes act on the same organ), the aggregate contact medium-specific risk ($AMR_{befghikt}(C_w)$) for contact medium i associated over all exposure routes from chemical e for cohort t associated with representative receptor of type h in exposure area g of site f from waste concentration C_w in WMU type b is given by the sum of the concurrent individual contact medium risk (defined by Equation (6)) for each exposure route:

$$AMR_{befghit}(C_w) = \sum_{k=1}^{nk} MR_{befghikt}(C_w) \quad \textbf{(8)}$$

  The contact medium-specific aggregate risk may be used as an indicator of the relative significance of the contact medium in conveying risks to the receptor.

  Similarly, the receptor-specific aggregate risk, $AR_{befght}(C_w)$, in the case of a carcinogen (or a non-carcinogen where the exposure routes act on the same organ), for a receptor cohort t associated with representative receptor of type h in exposure area g at time t from chemical e at site f and waste concentration C_w in WMU type b is given by the sum of the concurrent risks (ER) from each exposure route:

$$AR_{befght}(C_w) = \sum_{k=1}^{nk} ER_{befghkt}(C_w) \quad \textbf{(9)}$$

For carcinogens, if the receptor-specific aggregate risk exceeds a predetermined target risk or the maximum allowable risk threshold for the receptor, the receptor is said to be unprotected.

- Ecological risks are formulated in terms of a risk/hazard quotient type measure comparable to the human receptors. However, unlike human risk, ecological risk applies at the community and population level rather than at the individual receptor level.

  For example, the toxicity quotient for species that is exposed to constituent e, at site f, at time t is determined by

$$TQ_{befght} = \frac{CDI_{befght}}{CSCL_{eh}} \quad \textbf{(10)}$$

  where:

  TQ_{befght} = Toxicity quotient for chemical e, at site f, over exposure area g, for species h, at time t
  $CSCL_{eh}$ = Chemical stressor concentration limit for chemical e and species h (mg/L)
  CDI_{befght} = Chronic daily intake for chemical e, at site f, over exposure area g, for species h, at time t (mg/L)

  The chronic daily intake rate for species h that is exposed to chemical e, at site f, at time t, CDI_{befght}, is primarily a function of the following:

  Concentration of chemical in whole body prey (mg/kg)
  Daily quantity of prey ingested (kg/day)
  Fraction of contaminated material ingested
  Concentration of chemical in soil (mg/kg)
  Daily quantity of soil ingested (kg/day)
  Concentration of chemical in water (mg/L)
  Daily quantity of water ingested (L/day)
  Species-specific body weight (kg)

- Given that the impacts of different pathways (see Figures 3.1 and 3.2) can occur over significantly different time frames at a site and for a given individual receptor, all aggregations of doses and risks for a given cohort are carried out concurrently in time. Similarly all aggregations of protection measure statistics (e.g., number of receptors within a site that exceed a given target risk level) at the site are carried out concurrently in time.

  Figure 3.4 shows an example to illustrate how risks are aggregated concurrently in time. The example describes a case with two exposure pathways and one exposure route (ingestion of soil and ingestion of contaminated groundwater) for a representative receptor of type h for exposure area g at site f for a given waste concentration C_w of chemical e in a WMU of type b. The first two graphs show the pathway specific risks for each cohort t ($t=t_0,\ldots,T_{max}$) associated with the receptor. The last graph shows the exposure route (ingestion) specific risk for each cohort at the site that results from the concurrent aggregation of the individual pathway risks at the given exposure area.

3.2 ASSESSMENT STRATEGY

The HWIR99 assessment strategy represents the conceptual approach for applying the combination of models and data to develop national constituent-specific regulatory threshold levels for hazardous wastes. The assessment strategy for HWIR99 includes a "regional site-based" approach. The regional site-based

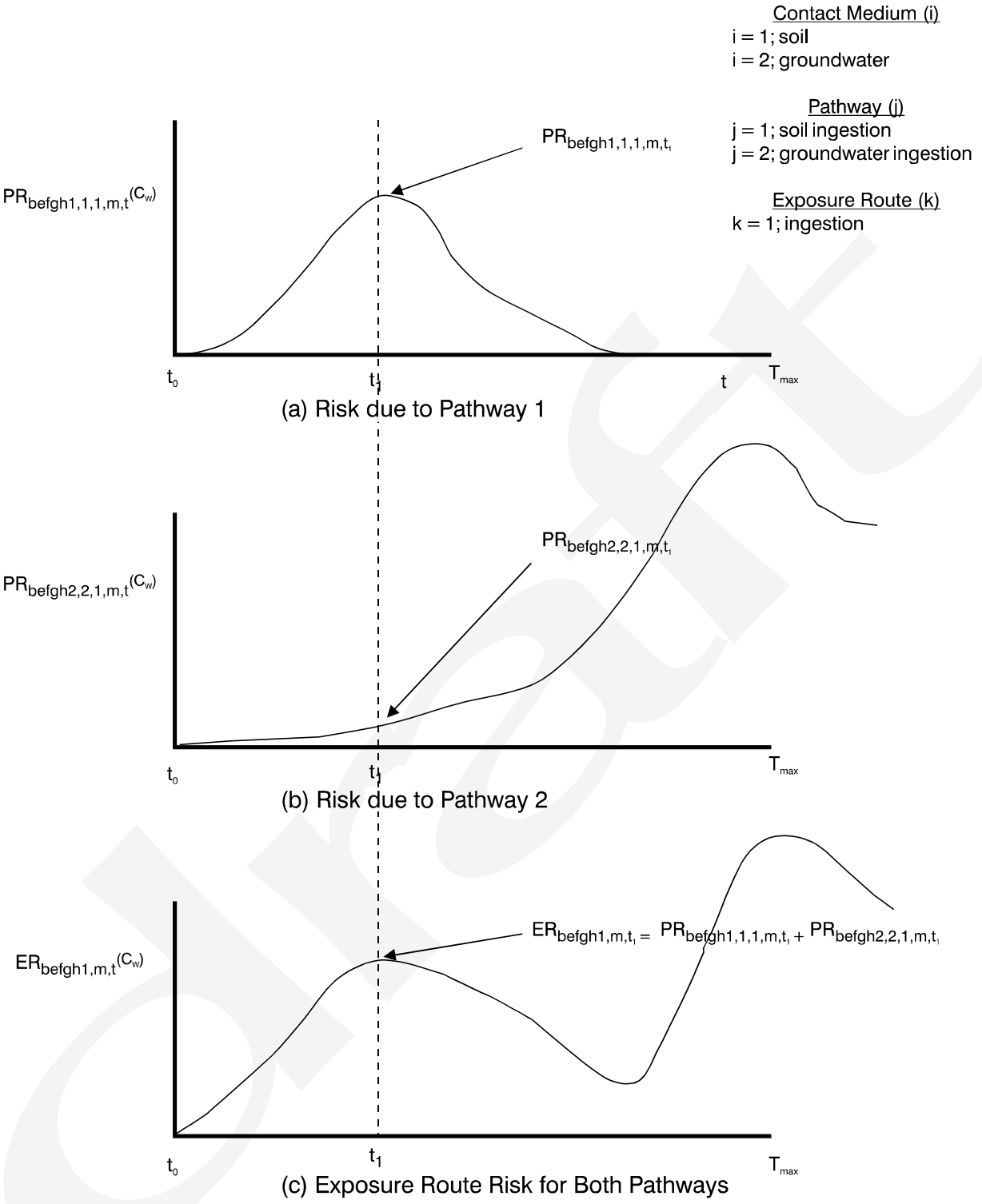

Figure 3.4 Pathway risks and exposure route risks for Site f.

approach was developed as part of EPA's Composite Model for leachate Migration with Transformation Products (EPACMTP) and was approved by EPA's Science Advisory Board (SAB, 1995).

The assessment approach for HWIR99 reflects an Agency decision to base exit levels on an assessment of potential health risks occurring at Subtitle D facilities (i.e., sites) where HWIR99 waste management units may be located. The objective here is to base the national constituent-specific exit criteria (i.e., allowable wastestream concentration for each HWIR constituent) on an assessment of risks under the widely varying environmental conditions, and receptor exposures associated with actual waste management units and locations.

The regional site-based approach embeds individual site-based assessments within a two-stage Monte Carlo simulation procedure. The overall objective of the iterative Monte Carlo procedure is to develop the nationwide distributions of risks and their uncertainty, as summarized by risk matrices, which can be queried to provide the basis for the development of the HWIR99 regulatory limits. The approach for describing the assessment strategy (Section 3.2.1) is to present first: the risk matrices, the intended output of the HWIR99 Technical Assessment; the protection measures which can be obtained by querying the risk matrices; and the regulatory framework, based on the proposed measures of protection, that establishes the procedure for determining the HWIR regulatory limits. Section 3.2.2 then describes the details of the Monte Carlo-based approach, including the general algorithm, and presents examples that illustrate the use of the proposed protection measures to establish the HWIR99 regulatory limits.

3.2.1 HWIR99 Strategy for Developing National Exemption Levels

The objective of the HWIR99 strategy for developing national exemption criteria is to develop a national database of site-based exposure and risk information, the risk matrices, that can be queried in different ways to support Agency decision makers in the establishment and implementation of exemption criteria. This section explains the contents of the database, the protection measures which define the different ways the database can be queried, and how the database/protection measures may be used to develop national exemption levels.

Section 3.2.1.1 defines the risk matrices that summarize the output of the HWIR99 technical assessment at each site. Section 3.2.1.2 presents the proposed protection measures that can be based on the selected protection criteria by querying the risk matrices at each site. Section 3.2.1.3 outlines the process for aggregating the protection measures of the individual sites to estimate nationwide impacts. Section 3.2.1.4 outlines the regulatory framework for establishing the HWIR99 regulatory limits based on the nationwide measures of protection presented in Section 3.2.1.3, and extends the approach to the case where the protection measures are characterized by uncertainty as well as variability. Finally, section 3.2.1.5 discusses alternative measures of protection.

3.2.1.1 Risk Matrices: HWIR Technical Assessment Output

The HWIR technical assessment output can be summarized through risk matrices that facilitate the process of developing exemption criteria. Four general risk matrix summaries are considered: pathway, contact medium, exposure route, and aggregate risk matrices.

Pathway Risk Matrix. For each site f, the baseline impacts for a given waste concentration C_w of chemical e in a WMU of type b can be summarized in a pathway risk matrix, $\underline{PR}_{bef}(C_w)$. The matrix consists of the pathway specific risks ($PR_{befghijkt}(C_w)$) for each pathway j associated with each contact medium i and exposure route k for each cohort t associated with representative receptor of type h at each exposure area

g. The pathway risks provide a baseline from which contact medium risks, exposure route risks, and aggregate risks can be computed for each cohort/receptor.

Contact Medium Risk Matrix. Contact medium risks can be summarized for each site by matrices for given values of the WMU type, chemical, and waste concentration. A contact medium matrix, $\underline{MR}_{bef}(C_w)$, consists of the contact medium specific risks ($MR_{befghikt}(C_w)$) for each contact medium i which are the respective sums of pathway specific risks from nj(i,k) pathways connecting contact medium i and exposure route k for each cohort t associated with each representative receptor type h at each exposure area g of site f for a given waste concentration (C_w) of chemical e in WMU type b.

Exposure Route Risk Matrix. Exposure route risks can be summarized for each site by matrices for given values of the WMU type, chemical, and waste concentration. An exposure route matrix, $\underline{ER}_{bef}(C_w)$, consists of the exposure route specific risks ($ER_{befghkt}(C_w)$) for each exposure route k for each cohort t associated with each representative receptor type h at each exposure area g of site f for a given waste concentration (C_w) of chemical e in WMU type b.

Aggregate Risk Matrix. An aggregate risk matrix, $\underline{AR}_{bef}(C_w)$, consists of the aggregate risks ($AR_{befghkt}(C_w)$) for each cohort t associated with each representative receptor type h at each exposure area g of site f for a given waste concentration (C_w) of chemical e in WMU type b.

3.2.1.2 Protection Measures for establishing HWIR99 Regulatory Exemption Levels

A protection measure provides a basis for establishing the HWIR99 regulatory limits. There are a number of alternative protection measures that can be considered in developing the exemption levels. Some of these are presented in Sections 3.2.1.4, and 3.2.1.5; others will be developed as the process of selecting exit levels progresses through the implementation stage. Note that the terms "protection measure" and "protection criterion" are used interchangeably in this document. In order to outline how regulatory exit levels will be developed, two candidate protection measures are presented in this section. The first proposed measure of protection is the nationwide distribution of risks for receptors of concern. Specifically, a regulatory limit is acceptable if the percent of nationwide receptors of concern that exceed a given risk level falls below an acceptable number. This protection measure can be applied, without loss of generality, to individual receptors, or combinations of receptors, as required by policy consideration. The second measure of protection is the nationwide distribution of sites that are protected. A regulatory limit is acceptable under the second protection measure if the percentage of protected sites nationwide is greater than a given target level.

3.2.1.2.1 Protection Measure based on Receptor Risk

The estimation of the number of receptors that exceed a given risk level (i.e., pathway specific risk, contact medium risk, exposure route risk, or aggregate risk) at a given site is calculated from the corresponding risk matrix. Since all receptors are being exposed from the same source, all inferences at the site level are based on concurrent year/exposure duration comparisons. Therefore, for a given site, all calculations are carried out individually for each concurrent year/cohort. The calculation consists of two steps. First, the number of concurrent cohorts that exceed the target risk level is determined for each exposure area; and second, the number of concurrent cohorts from different exposure areas that exceed the target level are added together. The result is the number of receptors that exceed the target level at the given site.

The remainder of this section will focus on aggregate (receptor-specific) risk, since the principles apply equally to the other types of risks, there is no loss in generality in limiting the discussion to the aggregate

risk case. More formally, let $RIND_{befght}(C_w, TR)$ represent an indicator that is set to 1 if the risk to cohort t associated with receptor m of type h in sector g of site f for waste concentration C_w of chemical e in WMU type b exceeds the target risk level TR; and is set to zero otherwise. Then in the case of human receptors, $RIND_{befght}(C_w, TR)$ is given by:

$$RIND_{befght}(C_w,TR) = 1, \quad if\ AR_{befght}(C_w) \geq TR; \\ = 0 \quad otherwise. \tag{11}$$

And in the case of ecological receptors, $RIND_{befght}(C_w, TR)$ is given by:

$$RIND_{befght}(C_w,TR) = 1, \quad if\ TQ_{befght}(C_w) \geq 1; \\ = 0\ otherwise. \tag{12}$$

where:

$TQ_{befght}(C_w)$ = Target toxicity quotient.

Then the number of receptors of type h in year t in site f that exceed the target risk, TR, for waste concentration C_w of chemical e in WMU type b is given by $NXR_{befht}(C_w,TR)$:

$$NXR_{befht}(C_w,TR) = \sum_{g=1}^{ng(f)} W_R(fgh)\ RIND_{befght}(C_w,TR) \tag{13}$$

where:

ng(f) = Number of exposure areas in site f

$W_R(fgh)$ = Weight for receptor type h in exposure area g in site f which is given by the number of receptors of type h in exposure area g in site f.

In the case of ecological receptors, information relating to population size of each representative receptor type h (species/community h) is not available, appropriate values for $W_R(fgh)$, other than unity, may be assigned to respective representative receptor types, to reflect the relative importance of the species.

3.2.1.2.2 Protection Measure based on Protected Sites

The implementation of this protection measure requires the definition of a protected site. Again, there are a number of alternative definitions. In general, a site can be defined as protective for a given receptor type h if the percentage of receptors of concern of a given type h that exceed a target risk level, TR, is less than or equal to an acceptable value q(h)%. In the most conservative case, q(h)% is set to zero for all h, so that a site is considered protective only if no receptors of the given type are exposed to risk levels above the target level.

More formally, let $SIND_{befht}(C_w, TR)$ represent an indicator that is set to 1 if site f is protective for cohort t associated with representative receptor of type h for waste concentration C_w of chemical e in WMU type b; and is set to zero otherwise. Then $SIND_{befht}(C_w, TR)$ is given by:

$$SIND_{befht}(C_w,TR) = 1, \quad if\ PXS_{befht}(C_w,TR) \le q(h)\%; \\ = 0\ otherwise. \qquad \textbf{(14)}$$

where:

$$PXS_{befht}(C_w,TR) = \frac{NXR_{befht}(C_w,TR)}{\sum_{g=1}^{ng(f)} W_R(fgh)} \times 100 \qquad \textbf{(15)}$$

with:

$W_R(fgh)$ = Number of receptors of type h in exposure area g at site f
ng(f) = Number of sectors in site f

Alternatively, the definition of a protected site can be extended to include all receptors, so that a site is protected if the percentage of receptors of concern of a given type h that exceed a target risk level, TR(h), is less than or equal to an acceptable value q(h)%, for all h.

More formally, let $ASIND_{beft}(C_w, TR)$ represent an indicator that is set to 1 if site f is protective for cohort/year t associated for waste concentration C_w of chemical e in WMU type b; and is set to zero otherwise. Then $ASIND_{beft}(C_w, TR)$ is given by:

$$ASIND_{befht}(C_w,TR) = 1, \quad if\ PXS_{befht}(C_w,TR) \le q(h)\%\ for\ all\ h; \\ = 0\ otherwise. \qquad \textbf{(16)}$$

As in the case of the site based protection measure, since all receptors are being exposed from the same source, all inferences at the site level are based on concurrent year/ exposure duration comparisons.

3.2.1.3 Nationwide Aggregation of Protection Measures

The previous section presented measures of protection that can be applied at a specific site. This section presents the method used to extrapolate the site specific results to a nationwide level.

The first step in determining the protection measures at the nationwide level is to repeat the site specific assessment described in the previous section to all facilities which have been selected in the sample design. The result, as shown in a column in the $N_f \times N_i$ matrix in Figure 3.5 would be a vector of size N_f; where each element (cell), corresponding to each of the N_f sites, consists of a risk matrix summarizing the corresponding risks for each receptor/cohort, in each exposure area for the corresponding site. This vector of risk matrices can then be queried to determine the protection measures at the nationwide level for each receptor type for any given waste concentration, chemical and WMU type.

For fixed: Chemical Type (e) Waste concentration (C_w) WMU Type (b)			UNCERTAINTY ITERATION				
VARIABILITY			1	2	3		N_i
	FACILITY	1	$MR_{b,e,1}(C_w, 1)$	$MR_{b,e,1}(C_w, 2)$			$MR_{b,e,1}(C_w, N_i)$
		2	$MR_{b,e,2}(C_w, 1)$	$MR_{b,e,2}(C_w, 2)$			$MR_{b,e,1}(C_w, N_i)$
		3					
						$MR_{b,e,f}(C_w, IT)$	
		N_f	$MR_{b,e,Nf}(C_w, 1)$	$MR_{b,e,Nf}(C_w, 2)$			$Mr_{b,e,Nf}(C_w, N_i)$

Figure 3.5 N_f x N_i Pathway Risk Matrix Output.

Note: Each element of the above matrix can be any risk matrix, e.g., $PR_{b,e,f}(C_w, IT)$, or $MR_{b,e,f}(C_w, IT)$, where $PR_{b,e,f}(C_w, IT)$ is the pathway risk matrix for WMU type b, chemical e, and site for waste concentration C_w and iteration IT, and $MR_{b,e,f}(C_w, IT)$ is the contact medium risk matrix for WMU type b, chemical e, and site for waste concentration C_w and iteration IT.

In general, if the number of receptors of type h in year t in site f that exceed the target risk, TR, for waste concentration C_w of chemical e in WMU type b is given by $NXR_{befht}(C_w,TR)$ as defined above, then the percentage of nationwide receptors of type h $PXR_{beht}(C_w,TR)$ over all sites that exceed the target risk, TR, is given by:

$$PXR_{beht}(C_w,TR) = \frac{\sum_{f=1}^{nf} W_S(f) \cdot NXR_{befht}(C_w,TR)}{\sum_{f=1}^{nf}\sum_{g=1}^{ng(f)} W_R(fgh) \cdot W_S(f)} \times 100 \quad \textbf{(17)}$$

where:
$W_S(f)$ = Sampling weight for site f
$W_R(fgh)$ = Number of receptors of type h in exposure area g at site f
ng(f) = Number of exposure areas in site f

The percentage of receptors that exceed a target risk level can also be calculated by combining all receptors. For example, the percentage of the nationwide total receptors, $APXR_{bet}(C_w,TR)$, that exceed the target risk, TR, is given by:

$$APXR_{bet}(C_w,TR) = \frac{\sum_{h=1}^{nh}\sum_{f=1}^{nf} W_S(f) \cdot NXR_{befht}(C_w,TR)}{\sum_{h=1}^{nh}\sum_{f=1}^{nf}\sum_{g=1}^{ng(f)} W_R(fgh) \cdot W_S(f)} \times 100 = \sum_{h=1}^{nh} PXR_{beht}(C_w,TR) \quad \textbf{(18)}$$

These equations apply equally to pathway, contact medium, exposure route and aggregate risk matrices.

Alternatively, the measure of the impacts can be described as the percentage of receptors of type h that are protected for the target risk level. Thus we can define the protection measure as the percentage of nationwide receptors $PPR_{bet}(C_w,TR)$ over all sites whose risk is below the target risk, TR. More formally, the nationwide percent protection for receptors of type h in year t for target risk, TR, for waste concentration C_w of chemical e in WMU type b is given by:

$$PPR_{beht}(C_w,TR) = 100\% - PXR_{beht}(C_w,TR) \quad \textbf{(19)}$$

Similarly, the nationwide percent protection for all receptors, $APPR_{beht}(C_w,TR)$, in year t for target risk, TR, for waste concentration C_w of chemical e in WMU type b is given by:

$$APPR_{bet}(C_w,TR) = 100\% - APXR_{bet}(C_w,TR) \quad \textbf{(20)}$$

In the case where the protection measure is based on the percentage of protected sites, then the analogous definition of protection is $PPS_{beht}(C_w,TR)$, the nationwide percentage of sites that are protected for receptors of type h in year t for target risk, TR, for waste concentration C_w of chemical e in WMU type b, which is given by:

$$PPS_{beht}(C_w,\ TR) \;=\; \frac{\sum_{f=1}^{nf} W_S(f) \cdot SIND_{befht}(C_w,\ TR)}{\sum_{f=1}^{nf} W_S(f)} \times 100 \qquad \textbf{(21)}$$

Similarly, the analogous definition of protection for all receptors is $APPS_{bet}(C_w,TR)$, the nationwide percent of sites that are protected for all receptors, in year t for target risk, TR, for waste concentration C_w of chemical e in WMU type b, which is given by:

$$APPS_{bet}(C_w,\ TR) \;=\; \frac{\sum_{f=1}^{nf} W_S(f) \cdot ASIND_{beft}(C_w,\ TR)}{\sum_{f=1}^{nf} W_S(f)} \times 100 \qquad \textbf{(22)}$$

3.2.1.4 Regulatory Scheme

The previous section outlined a procedure for deriving an estimate of the nationwide impacts to receptors of concern. The impacts are defined by a "protection measure" based on either the percentage of receptors that are below the target risk level or the number of protected sites. A defined protection measure is determined for each site and exposure area by querying the relevant risk matrices that provide the raw risk data for all receptors. Depending on the protection measure and the specific constituent, the relevant risk matrices can include pathway specific, contact medium specific, exposure route specific, and aggregate risk specific matrices for a single type of receptor, for groups of selected receptor types, or for all receptor types.

In general, the derivation of a regulatory limit for a given chemical consists of two steps. First, derive for each WMU type a waste concentration limit that satisfies the protection measure criteria for the given WMU type; and second, set one or more regulatory limit (exit criteria) from the WMU specific concentration limits on the basis of policy decisions.

The remainder of the discussion is based on the receptor based protection measure for the criteria based on all receptors. The procedure outlined below is also applicable to the site based protection measure. Extension of the discussion to the site based protection measure would only require the replacement of every instance of $APPR_{be}(C_w,TR)$ with $APPS_{be}(C_w,TR)$. The extension to the receptor type specific case would require a similar replacement of the applicable notation.

For a given WMU type, the regulatory waste concentration is selected as the largest waste concentration that meets the protection measure criteria. For the purposes of this discussion, the protection measure criteria are met if at least p% of the nationwide receptors have risk below the target risk for every concurrent set of cohorts. For the given protection measure, the relevant percent protection, $APPR_{be}(C_w,TR)$, for a given chemical and WMU type, occurs in the year with the minimum level of protection, i.e.:

$$APPR_{be}(C_w,TR) = MIN_t \ (APPR_{bet}(C_w,TR) \mid t_0 \leq t \leq T_{\max}) \quad \textbf{(23)}$$

Focusing on the year with the minimum level of protection guarantees that every concurrent set of nationwide cohorts meets the protection measure. Given the protection criteria, a concentration waste limit, $C_{w,limit,b,e}$, for chemical e is selected as the regulatory limit for a given WMU type b, if $C_{w,limit,b,e}$ is the largest waste concentration such that:

$$APPR_{be}(C_{w,limit,b,e}, TR) \geq p\%. \quad \textbf{(24)}$$

Once the limits for each WMU type are determined, one or more regulatory concentration waste limits $C_{w,limit,e}$ for chemical e are selected from the WMU specific limits based on policy considerations.

The protection measures, however, are characterized by uncertainty. In the presence of uncertainty, the protection measure is modified to include the additional criterion that the percent protection must be met with at least a specified level of confidence. An example of this modified protection measure is that 90% of the nationwide receptors of concern would be exposed to risks less than 10^{-6} with at least a 95% level of confidence (probability).

For the uncertainty case, the output data base used to derive the protection measures consists, as shown in Figure 3.5, of a matrix rather than a vector of risk matrices. In effect, the $N_f x N_i$ output risk matrix consists of N_i iterations of the single vector of risk matrices presented in a column in Figure 3.5; where each column represents an alternative realization of the risk matrices resulting from the uncertainty in the characteristics that describe a given simulation scenario.

For a given chemical waste concentration, each column ($IT = 1,\ldots,N_i$) of the output data base can be queried separately to determine IT different values of the minimum nationwide percentage of receptors that are protected for a given target risk level, $APPR_{be}(C_w,TR,IT)$ for a given chemical and WMU. This effectively results in N_i separate estimates of $APPR_{be}(C_w, TR, IT)$, $IT = 1,2,\ldots,N_i$ that reflect the uncertainty in their prediction. Together, the N_i iterations of APPR can be used to establish confidence levels (or probability values in a Bayesian context) that the given protection measure will be met. A description of the Monte Carlo algorithm used to generate the $N_f x N_i$ output risk matrix, as well as examples describing how the nationwide exit levels are derived are presented in section 3.2.2.

This discussion applies to both human and ecological receptors. Each can be addressed with the proposed framework, but must be addressed separately. In the case of humans, the primary protection measures will involve the nationwide percentage of protected individual receptors; while in the case of ecological receptors, the primary measures involve the nationwide percentage of receptor species/communities. The measures are not directly comparable. Therefore, a separate regulatory limit is derived for humans and for ecological receptors. The final limit(s) is (are) given by the more restrictive of the two.

There are other alternative measures of protection that could be used to derive regulatory limits. The protection measure based on the percentage of all receptors protected at a given aggregate target risk level provides a convenient starting point for presenting the methodology. One alternative is to focus on the number of protected sites rather than receptors. A discussion of other alternative measures of protection is presented in the next section, Section 3.2.1.5.

3.2.1.5 Alternative Measures of Protection

The previous section outlined a regulatory framework based on two protection measures. The first is a function of the percent of nationwide receptors that are exposed to risks less than a given target risk level. The second is a function of the percentage of nationwide that are protected. There are a number of other alternative definitions of protection which could be queried from the output database. One alternative measure is a variation on the receptor based percent protection measure that involves both primary and secondary criteria. The primary criteria would be met if at least p% of all of the receptors have risk below the target risk level; and the secondary criteria would be met if no less than q% of any given type of receptor have risk below the target risk level with $p\% > q\%$. Both criteria would have to be met with a minimum level of confidence in order to satisfy the overall protection criteria.

Another alternative involving primary and secondary criteria would be to use the same primary criteria, but select the secondary criteria to include a separate criteria for the average risk (or some other statistical measure) of all receptors, or subsets of receptors, that exceed the primary criteria. Thus for example, a given waste concentration would meet this protection measure criterion if at least p% of all receptors had risk less than the primary target risk, and the average risk of the receptors, for receptors that exceeded the primary target risk, is below a secondary target risk level. Again, both criteria would have to be met with a minimum level of confidence in order to satisfy the overall protection criteria.

There are numerous other possibilities that could be queried from the output risk matrices, including variations on the site-based protection measure, and variations on the concurrent cohort requirements. In particular the regulatory framework presented in Section 3.2.1.4 for both the receptor risk and site based protection measures is based on concurrent cohorts both within a site as well as between sites. An alternative is to develop the regulatory framework so that the requirement for concurrent cohorts within a site is maintained as discussed in Sections 3.2.1.2.1 and 3.2.1.2.2, but does not require concurrent cohorts between sites.

Ultimately, the criteria will take the form that a regulatory waste concentration is selected if it meets the adopted measure of protection with a given level of confidence. The proposed two-stage Monte Carlo framework is sufficiently general to accommodate these options.

3.2.2 Monte-Carlo Approach

This section presents a general outline of the Monte-Carlo approach proposed for the production of the $N_f \times N_i$ output matrix that forms the basis for the regulatory framework outlined in Section 3.2.1.2. The remainder of this section is organized as follows. The objectives of the Monte Carlo procedure are presented in Section 3.2.2.1. The proposed Monte Carlo implementation strategy is presented in 3.2.2.2. The latter section includes a general outline of the proposed Monte Carlo method, together with sample queries and outputs.

3.2.2.1 Monte-Carlo Objectives

The proposed Monte-Carlo procedure is designed to meet the following objectives:

- 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.

3.2.2.2 Monte-Carlo Implementation Strategy

3.2.2.2.1 Ideal Conditions

The validity of a Monte Carlo implementation depends ultimately on the amount, type, and quality of the data available to estimate the probability distributions of the Monte Carlo inputs. The fundamental question that the proposed framework is designed to answer for a given chemical can be stated in the following way: If a "receptor of concern" is defined as all receptors of a given type that currently reside within a specified radius of all currently existing Subtitle D waste management facilities in the continental U.S., then what percent of the total number of current receptors of concern would be exposed to risk/hazard quotient levels above specified target levels if each facility were to manage the chemical at the same concentration at all facilities.

Clearly, any attempt to determine nationwide risks is a challenge. Performing a risk assessment at a site-specific level is difficult enough. Extending the site-specific effort to a nationwide scale introduces an additional, and significant, layer of difficulty. Ideally the nationwide risk assessment would be performed in four steps. First, identify all current Subtitle D waste management facilities in the continental U.S. Second, collect all of the site-specific data necessary to characterize each facility and associated site/receptor characteristics, and relevant processes. Third, develop a site-specific mathematical model to predict the impacts at each site; and fourth, run the site-specific model at each of the sites and aggregate risks to predict the nationwide impacts to the "receptors of concern".

Under ideal conditions, the HWIR99 Monte Carlo approach would be based on the following database:

1) A statistically designed sample of waste management units from the target population of 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 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.

This ideal data set, together with sufficient computational resources, provides a solid foundation for the identification and estimation of the relative magnitude of applicable sources of uncertainty (e.g., sampling errors, data errors, model prediction errors, non-sampling errors) and variability; and the development of a two-stage Monte Carlo algorithm that incorporates and separates the effects of uncertainty and variability. This separation allows for the estimation of uncertainties associated with given measures of variability, which form the basis of the regulatory framework presented in Sections 3.2.1.2. For an introduction to the topic of uncertainty and variability in HWIR99, the reader is referred to Appendix A. The appendix

provides a summary of the various sources of uncertainty and variability, and a discussion of the importance of separating uncertainty and variability.

3.2.2.2.2 Limitations in Implementation of the Monte Carlo Approach

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. 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 to receptors.

Second, the development of a site-specific model for each facility is impractical. This implies that a generic model will need to be developed that can be applied at all sites. A generic model is generally less able to approximate causal relationships than a site-specific model. Additionally, scheduling 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 the 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.

Fourth, resource constraints 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 data bases, which raises the question of the representativeness of the data to the target population. Examples of parameters that cannot be practically obtained at the site specific level for all sites include receptor exposure/response physiologic 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 errors 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 estimated measures of variability obtained from a two stage Monte Carlo will only reflect the identified sources of uncertainty for part of the variability. The unestimated 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 eliminate the sources of variability and uncertainty that have the least impact while meeting the budgetary, scheduling, and computational capacity constraints imposed on the problem.

3.2.2.2.3 HWIR99 Site Based Approach Monte Carlo

This section presents a proposed Monte Carlo structure to support the regulatory framework outlined in section 3.2.1.4. Although this data base is a departure from the ideal situation discussed above, 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 data base include:

1) A probability subsample of 201 WMU facilities from a stratified sample national survey of WMU facilities (U.S. EPA, 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 meteorologic 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

The Monte Carlo algorithm will follow the general form of the two stage Monte Carlo presented in Section 3.2.2.2.3.1. 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. 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 (1997).

3.2.2.2.3.1 *General Monte Carlo Algorithm*

A general and idealized form for the HWIR99 Monte Carlo approach is presented in the flowchart in Figure 3.6. The flowchart is designed to illustrate and help explain the general steps of the approach for the primary protection measure described in previous sections. Specific details of the algorithm are not included since these will depend on the protection measure adopted, the type and amount of data available

to estimate the needed probability distributions, the sources of variability and uncertainty that are significant, correlations among parameters, the methods used to estimate and model the various sources of variability and uncertainty, and computational efficiency considerations.

In its present form, the algorithm will produces an output $N_f x N_i$ matrix that can be queried, as described in Section 3.2.2.2.3.2, to determine whether a given waste concentration meets the protection measure criteria within a given level of confidence. For the purpose of this illustration, each cell of the $N_f x N_i$ matrix, $MR_{bef}(C_w,IT)$ corresponds to the contact medium risk for a given waste concentration, chemical, WMU type, site, and iteration (IT). Alternatively, the algorithm could also have been written so that each cell corresponds to the pathway specific risk. In practice, since the storage of the risk matrices at the pathway risk level may impose excessive computational requirements, the output database will likely be based on the contact medium risk matrices. The algorithm is sufficiently general that the basic elements apply whether interest is on the pathway or contact medium matrices.

Each row of the matrix corresponds to a sampled facility; and each column represents an alternative realization of the risk matrices resulting from the uncertainty that characterize a given simulation scenario. The level of confidence is derived by determining the protection measure independently for each iteration. The resulting N_i estimates of the protection measure represent a conditional distribution that allows the estimation of the probability (confidence level) that a given measure of protection will be met for a given waste concentration, C_w.

At this stage of development, the conditional distribution represents the uncertainty in the protection level only due to sampling error for a given value of C_w. It does not address data measurement errors, or model prediction component errors. Additionally, it does not address the more general case that includes uncertainties due to misspecification of the probability distribution functions (pdf) and parameters used to describe the different uncertainties, misspecification of the assumed pdf models that describe variability, errors associated with non-probability samples, or sampling of non-target populations. Such sources of uncertainty can be included or at least evaluated, at least initially through subjective measures, but are not addressed in this example. Ultimately, the decision on whether to incorporate a source of uncertainty will depend on the results of sensitivity testing to determine the significance of each source relative to the selection of the regulatory waste concentrations.

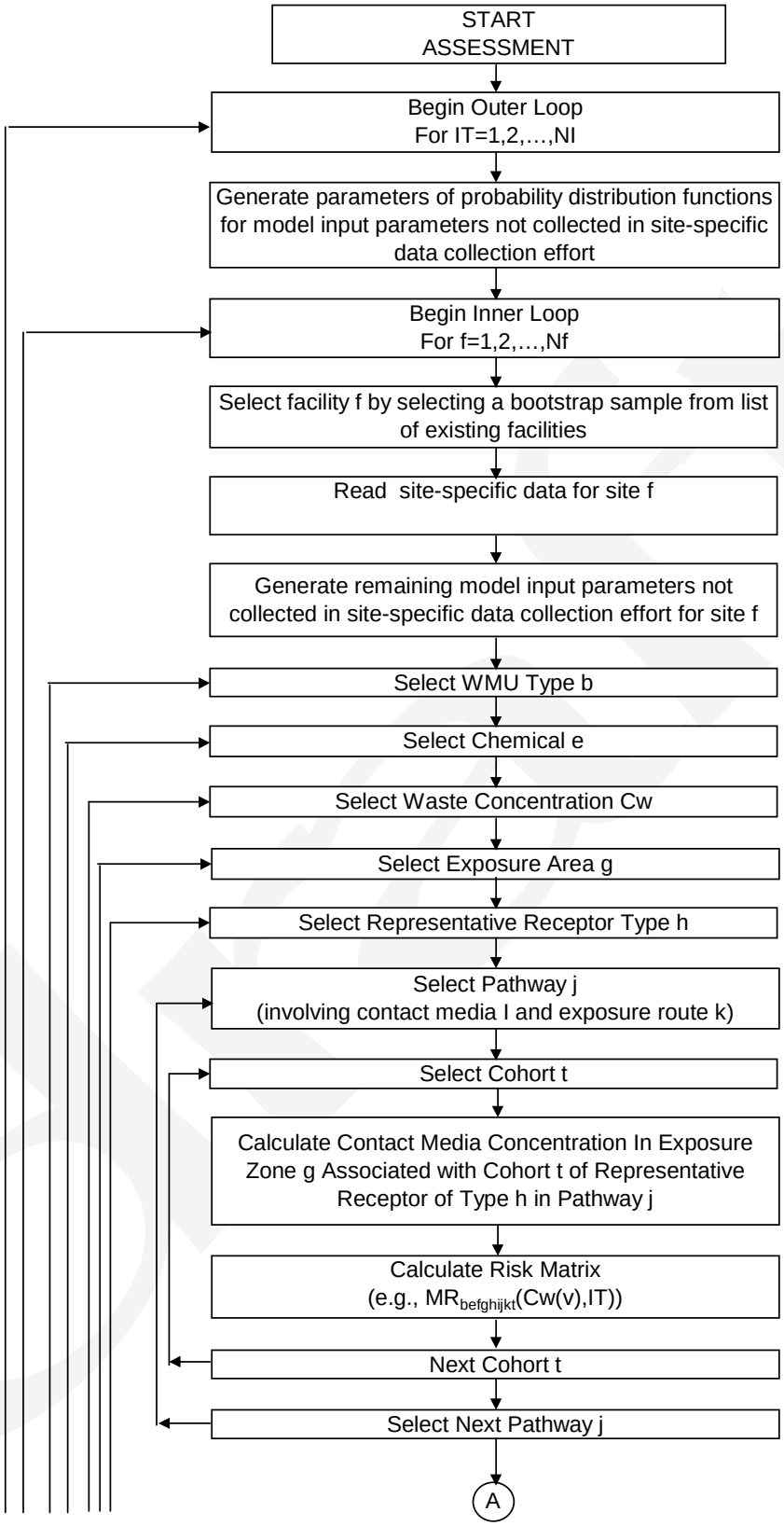

Figure 3.6 General Assessment Flowchart.

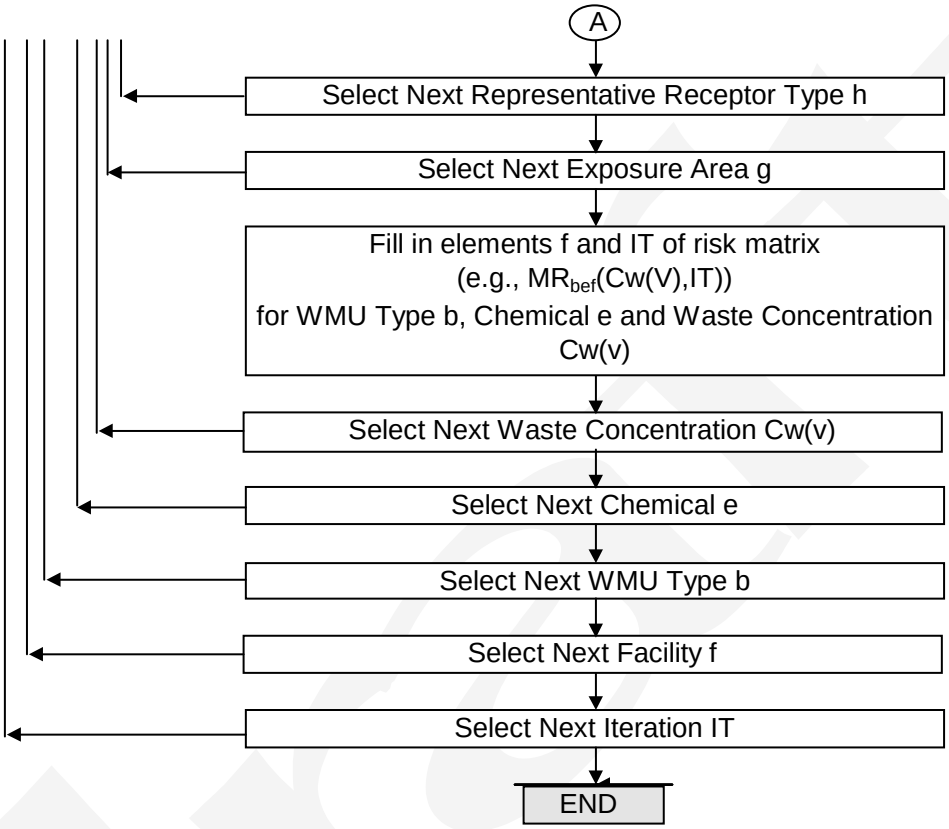

Figure 3.6 (continued) General Assessment Flowchart.

For this example, sampling error uncertainty for input parameters not measured directly at each site is incorporated through a Bayesian and/or parametric bootstrap approach. In the actual case, a combination of parametric and nonparametric methods are possible. The specific form that will be adopted will depend on the type and amount of data available to estimate the needed probability distributions.

Sampling error uncertainty in input parameters that are measured directly at each site for this example is simulated in the algorithm by a nonparameteric bootstrap of the facility within the inside loop. Each time the facility is selected in a bootstrap sample, all of the site-specific information measured at the facility is included in the sample.

The algorithm starts at the outer loop by generating and storing the parameters of the probability distribution functions that describe the between site variability of the various model input parameters that are not collected at each site. That is, those parameters whose between site variability will be based on regional and/or national distributions. The probability distribution function parameters are generated from probability distributions that reflect the uncertainties in their estimation. Examples of input parameters that will not be collected at the site-specific level include climatic, hydrogeologic, ambient water quality, physiologic and behavioral receptor exposure factors, and chemical specific characteristics.

Note that the pdf parameters generated in the outer loop remain fixed for all facilities for a given iteration. For example, suppose the pdf that describes the between facility variability of groundwater temperature at the national scale is normally distributed with some mean and variance. Then in order to generate the groundwater temperature for each of the N_f facilities in the given iteration, the mean and variance of the groundwater temperature pdf are generated in the outer loop. Assuming that the mean and variance generated in the outer loop are 20°C and 40 $(°C)^2$, respectively, the groundwater temperatures for all N_f facilities in the given iteration are generated from a normal distribution with mean of 20 °C and variance of 40 $(°C)^2$. The pdf for the groundwater temperature in the next iteration would have a different mean and variance which reflect the uncertainty in the parameters due to sampling error. In the case of parameters that are characterized by regional probability distributions, the pdf parameters will vary between regions within an iteration, but the pdf parameters of a given region will remain constant for all sites within a region within an iteration.

The inner loop begins after the generating in the outer loop the parameters of the pdfs that describe the between facility variability of the input parameters that are not measured at the site. Facilities in the sample were randomly selected to represent the nation-wide population of industrial facilities that generate Subtitle D non-hazardous waste and handle it on-site. Bootstrap sampling/analysis can be used. Weights can be incorporated into this process which account for several known sources: sampling with replacement from finite population; and size of facilities.

The next step in the algorithm involves generating the remaining input parameters that were not measured at the site but are needed to describe a simulation scenario for the given site. These parameters are generated using the corresponding pdfs, conditional on the fixed pdf parameters generated in the outer loop. The conditional pdfs reflect variability of the parameters between sites, between sectors at a site, and within a site as applicable, as well as any relevant correlations between parameters.

Once the site/facility scenario is generated, the next twenty steps of the algorithm involve the calculation of the pathway risk matrices, $PR_{bef}(C_w(v), IT)$ for every pathway, or contact medium risk matrices, $MR_{bef}(C_w(v), IT)$ for every contact medium, for the representative receptor of every receptor type at the site for every chemical, WMU type and waste concentration. There are a number of intermediate steps involved in the calculation of the risk matrices that are not shown in the algorithm. The first of these steps involve using the input parameters generated for the facility to calculate the exposure zone concentrations

for each contact medium associated with each pathway and each cohort of each representative receptor of each type at the site for each chemical, WMU type and waste concentration.

The next step involves using the exposure factors (e.g., exposure duration) generated in the second step for each representative receptor of each type at the site to calculate the risk matrices for each cohort of each representative receptor of each type, for each chemical, WMU type and waste concentration. The calculation of risk, by policy decision, does not incorporate model error.

These steps are repeated for all N_f selected sites to calculate the N_f set of risk matrices for all sites in the given iteration. The outer loop is then repeated N_i times to produce the $N_f x N_i$ sets of risk matrices that provide the database that is queried in the next section of the algorithm.

3.2.2.2.3.2 *Output Queries*

In general, for any measure of protection, each of the N_i columns in the $N_f x N_i$ matrix can be queried to produce one estimate of the protection measure. Together, the N_i estimates of the protection measure can be used to create a probability distribution that a) describes the uncertainty in the protective measure; and b) provides an estimate of the probability (uncertainty) that the protection measure will be met.

The rest of the algorithm, as presented in Figure 3.7 illustrates how the querying of the $N_f x N_i$ sets of matrices can be used to select a regulatory limit for each chemical in the case where the protection measure is the nationwide percentage of all receptors that are protected for a given target risk level.

The query process is initiated by specifying a trial waste concentration, C_{wbe}, for a given chemical and WMU type. The first step involves calculating from the corresponding sets of N_f pathway risk matrices, $PR_{bef}(C_{wbe},IT)$, $f=1,\ldots,N_f$, for the given chemical, waste concentration and WMU type b, the nationwide percentage of receptors that are protected at the target level risk TR for each cohort t in iteration IT, $APPR_{bet}(C_{wbe},TR,IT)$. Note that if C_{wbe} was not specifically included as one of the waste concentrations used in calculation of the $N_f x N_i$ matrices in the first part of the algorithm, the nationwide percent protection can be estimated by interpolating values of the matrices corresponding to waste concentrations that bound C_{wbe}.

The nationwide percentage of receptors that are protected at the target level risk TR for the given waste concentration C_{wbe}, WMU type b and chemical e, for iteration IT, $APPR_{be}(C_{wbe},TR,IT)$, is then calculated by selecting the concurrent cohort/year t that gives the minimum percent protection:

$$APPR_{be}(C_{wbe},TR,IT) = MIN_t \left(APPR_{bet}(C_{wbe},TR,IT) \mid t_0 \preceq t \preceq T_{max}\right)$$

Repeating the process for all iterations, gives N_i values of the measure of protection, $APPR_{be}(C_{wbe},TR,IT)$, $IT=1,2,\ldots,N_i$ that can then be used to estimate whether the trial waste concentration, C_{wbe}, meets the percent protection criteria for the given WMU type and chemical with a sufficiently high probability.

If the trial waste concentration does not meet the protection criteria, a new waste concentration is tried until the largest concentration that meets the protection criteria is found. Although not explicitly addressed in the flowchart, it should be noted that the process of selecting alternate waste concentrations can be optimized by using efficient search techniques.

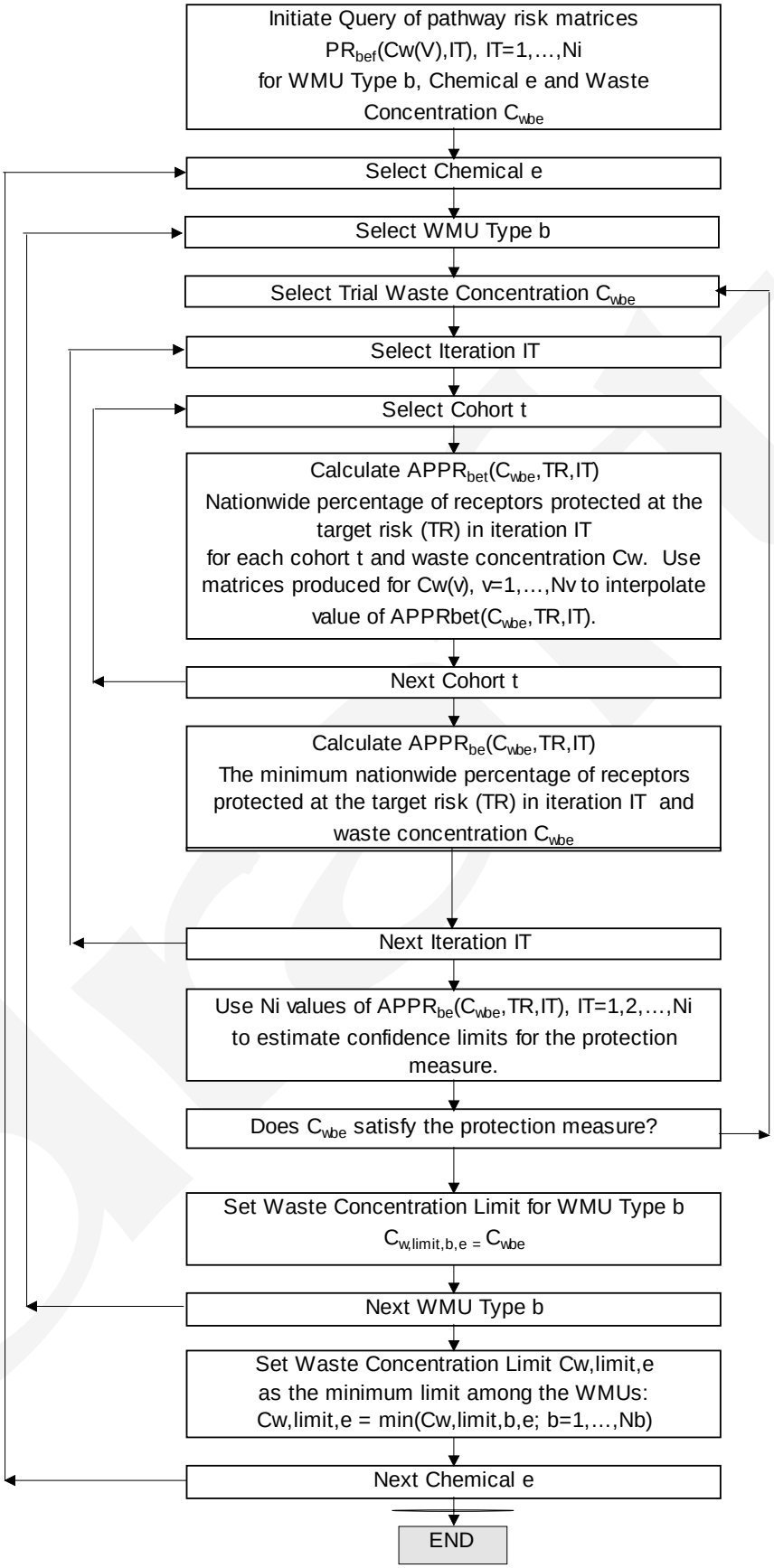

Figure 3.7 Exit Level Determination Flowchart.

The process is repeated for each WMU type for the given chemical to yield the largest waste concentration limit, $C_{w,limit,b,e}$ for each WMU type ($b=1,\ldots,N_b$) that meets the protection criteria. One or more regulatory limit for the given chemical, $C_{w,limit,e}$, is selected among the limits established for each WMU type on the basis of policy considerations. The process is repeated for each chemical to obtain a regulatory waste concentration limit for each chemical.

3.2.2.2.3.3 Example Monte Carlo Output

This section presents examples of output that can be obtained by querying the data base generated by the two-stage Monte Carlo algorithm discussed in the previous section. Figure 3.8(a) presents an example corresponding to a query for a target risk level of 10^{-6} from the N_i (columns) iterations of risk matrices corresponding to a waste concentration of 10^{-3} mg/kg. The figure indicates that there is a 5% chance that the level of protection (% of receptors that would be protected at the target risk level for the given waste concentration) would be less than or equal to 85%. Similarly, there is a 25% chance that less than or equal to 93% of the receptors would be protected at the target risk level for the given waste concentration.

The result of repeating the query for different target risk levels for the same waste concentration 10^{-3} mg/kg is illustrated by Figure 3.8(b), which presents the uncertainty in the percent of protected receptors for each risk level. From Figure 3.8(b), it can be inferred that there is a 95% chance that setting the waste concentration regulatory limit to 0.001 mg/kg, would result in at least 85% of the receptors protected to a 1E-6 risk level (*or 5% chance that, at the risk level of 1E-6, less than 85% of the receptors will be protected*), and at least 90% of the receptors protected to a 1E-5 risk level. Similarly, there would be a 95% chance that at least 95% of the receptors would be protected to the 1E-4 risk level, and at least 50% of the receptors would be protected to the 1E-7 risk level.

Querying the output data base for different waste concentrations can produce the set of graphs such as those shown in Figures 3.9 (a), 3.9 (b), and 3.9 (c). The figure shows how the percent protection varies as a function of the target risk, the waste concentration and the confidence limit; and can be used to select the waste concentration that meets a specified protection measure. These types of figures could also be produced for subsets of receptors to investigate the effects of selecting a waste concentration on secondary protection measures.

In particular, if the exit level criteria requires that at least 85% of the receptors must be protected at the 1E-6 risk with at least a 95% confidence level, then the exit level would be 0.001 mg/kg (Figure 3.9 (a)). If on the other hand, if the exit level criteria requires that at least 90% of the receptors must be protected at the 1E-6 risk with at least a 95% confidence level, then the exit level would have to be less than 0.001 mg/kg. In this case, Figure 3.9(a), would be used to determine the appropriate exit level.

As is evident by the figures presented in this section, the most notable effect of introducing uncertainty in the estimation of the protection measures is that the regulatory criterion for accepting a waste concentration limit must be modified to incorporate a minimum probability that the protection level will be obtained.

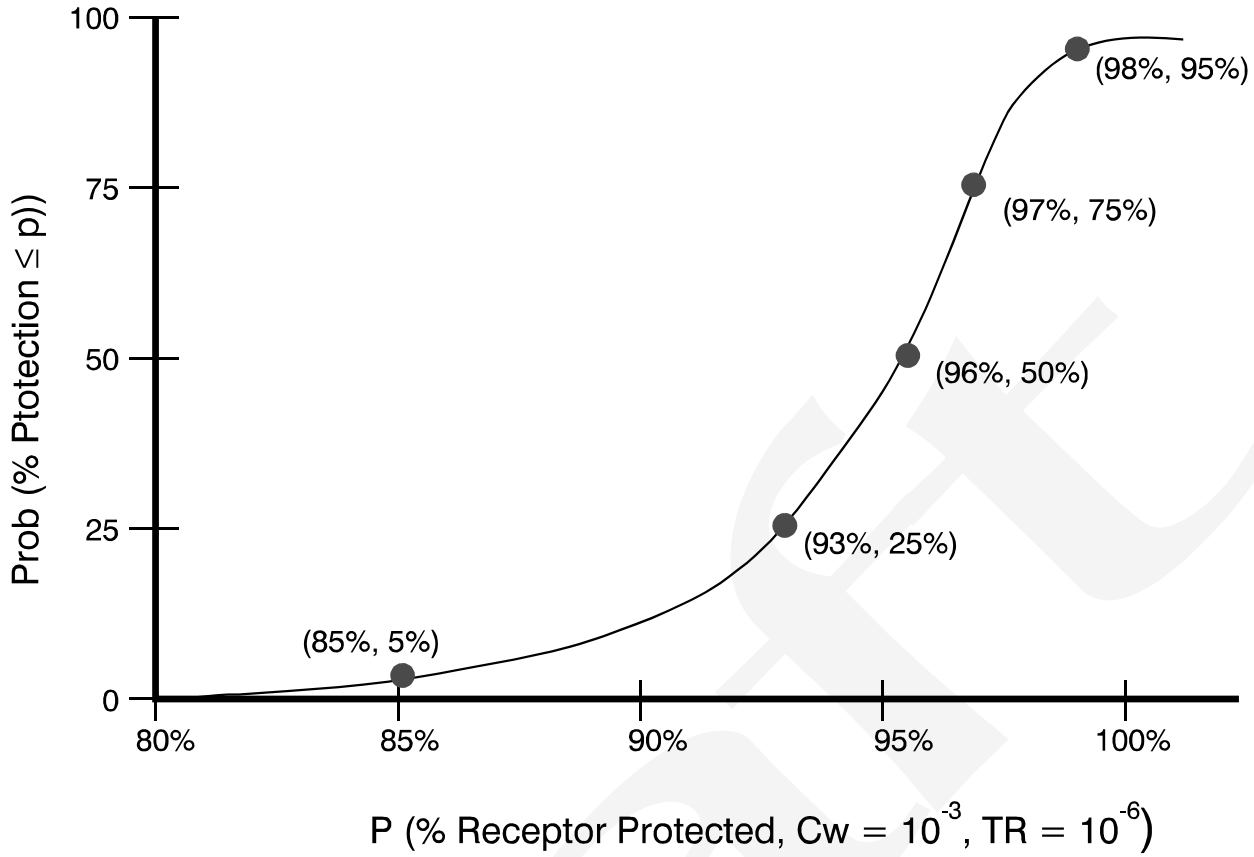

Figure 3.8a **Probability that percent protection is less than P for a given waste concentration and target risk level.**

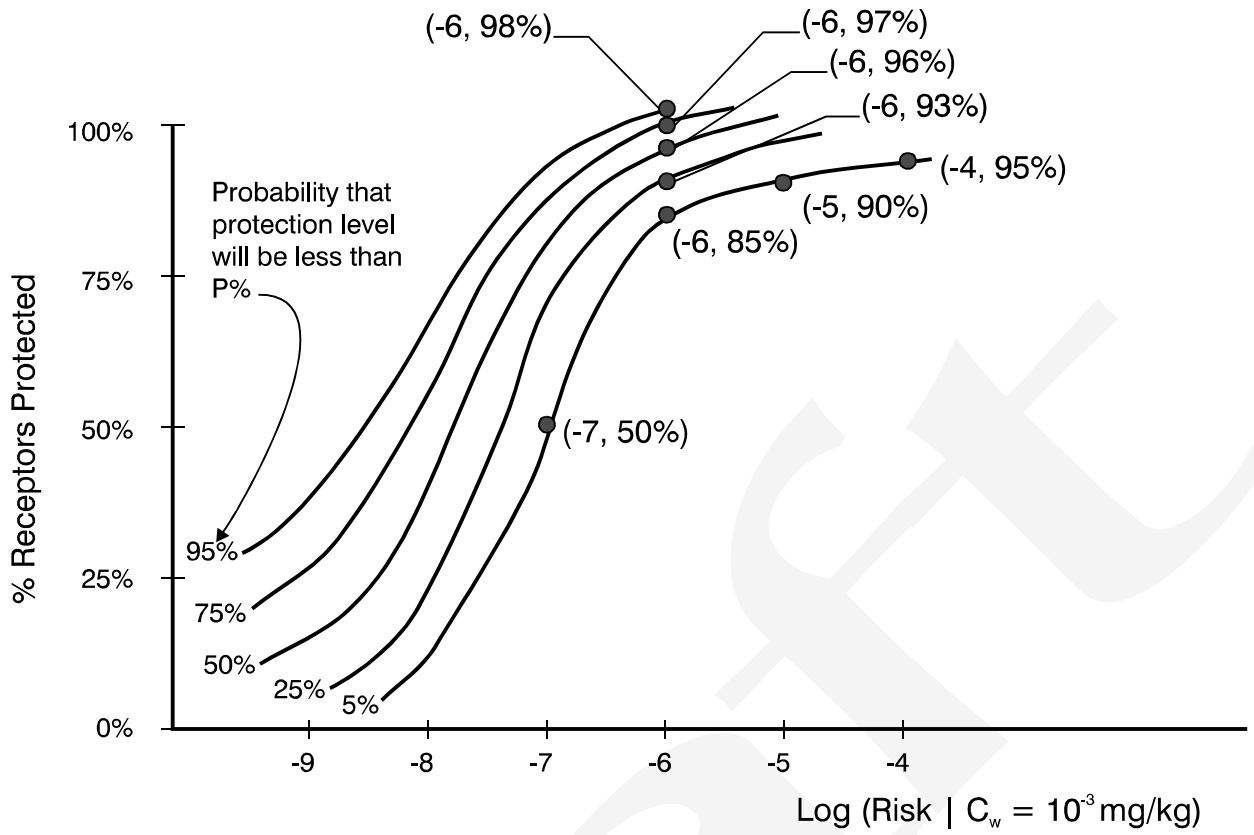

Figure 3.8b **Percent of receptors protected for different risk levels and $C_w = 10^{-3}$ for N_i Monte-Carlo iterations.**

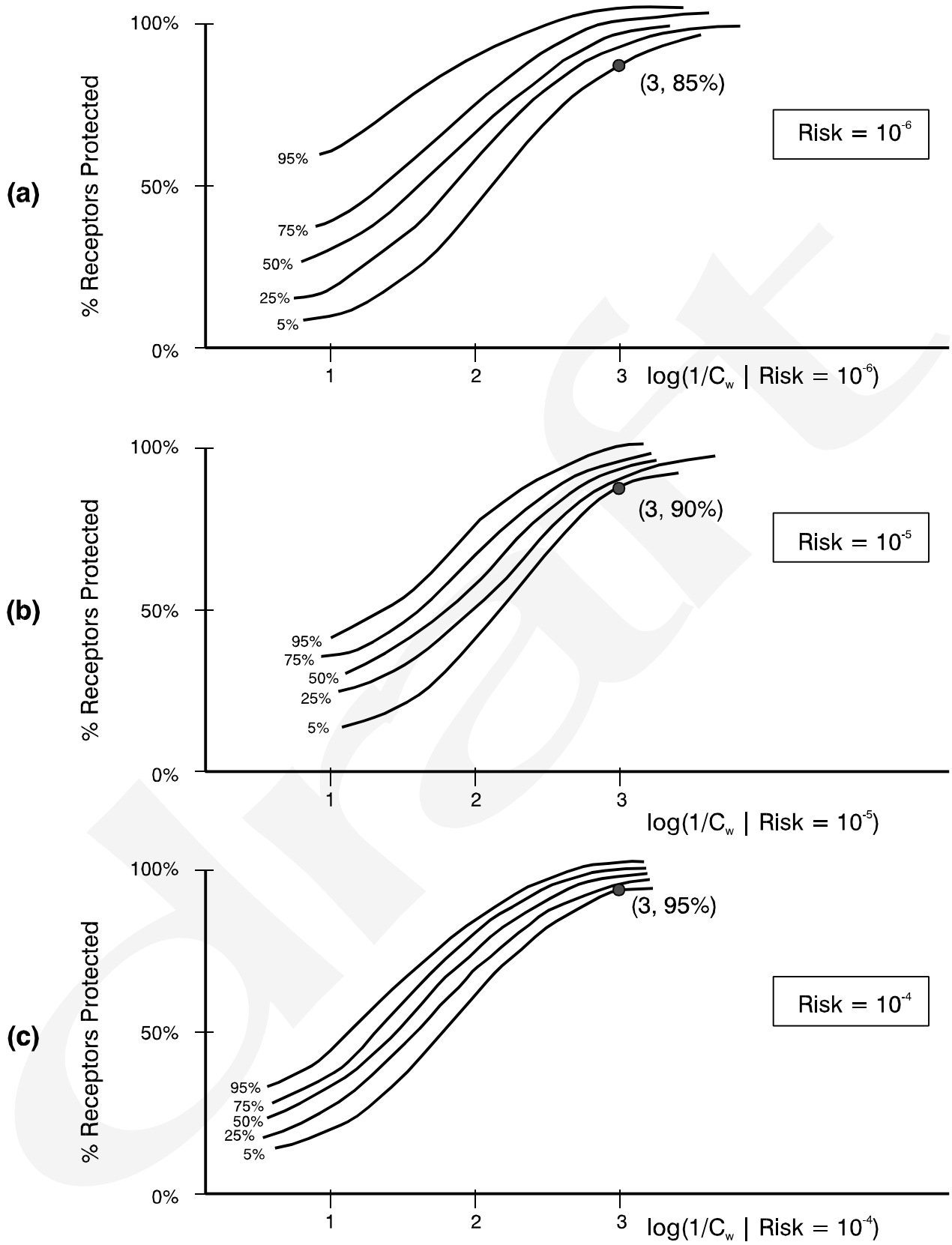

Figure 3.9 **Percent of receptors protected for different waste concentrations and risk levels**.

3.2.3 AN EXAMPLE OF 3MRA IMPLEMENTATION

As stated previously that there are a number of alternative measures of protection that can be used in the 3MRA framework, an example, based on one of the possible measures of protection, is given in this section to help elucidate the risk assessment methodology presented earlier. This example is based on the percentage (p) of sites that are protective of both human and ecological receptors. In this example, a site is said to be protective when at least a given percentage of the receptors (say 99 percent) are exposed to risks smaller than the threshold risk.

Based on the measure of protection adopted for this example, regulatory standards are based on the calculation of the conditional probability distribution, $f(C_w|p)$, of constituent-specific regulatory waste protection concentration, C_w, for a given percentage, p percent, of sites that are protective. The distribution reflects the uncertainty in C_w for a given p, or vice versa, due to input parameter measurement and sampling errors, and model prediction errors. The family of conditional distributions can be used to generate isopleths of the probabilities that a given protection level will be greater than the stated value for the given protective waste concentration, C_w. An illustration is given in Figure 3.10.

A regulatory level of the protective waste concentration is selected as the concentration value that results in the protection of at least p% of the sites with at least q% probability that a given protection level will be above the stated value. In Figure 3.10, for example, setting the regulatory waste concentration to 0.001 ($\log (1/C_w) = 3$), would result in the protection of at least 85% of the sites with at least 95% probability that protection level would be equal to or exceed the stated value of 85% (or alternatively, a 5% chance that the protection level would be less than 85%). In general, it would be recommended that the regulatory scheme lead to more conservative regulatory levels as the uncertainty increases.

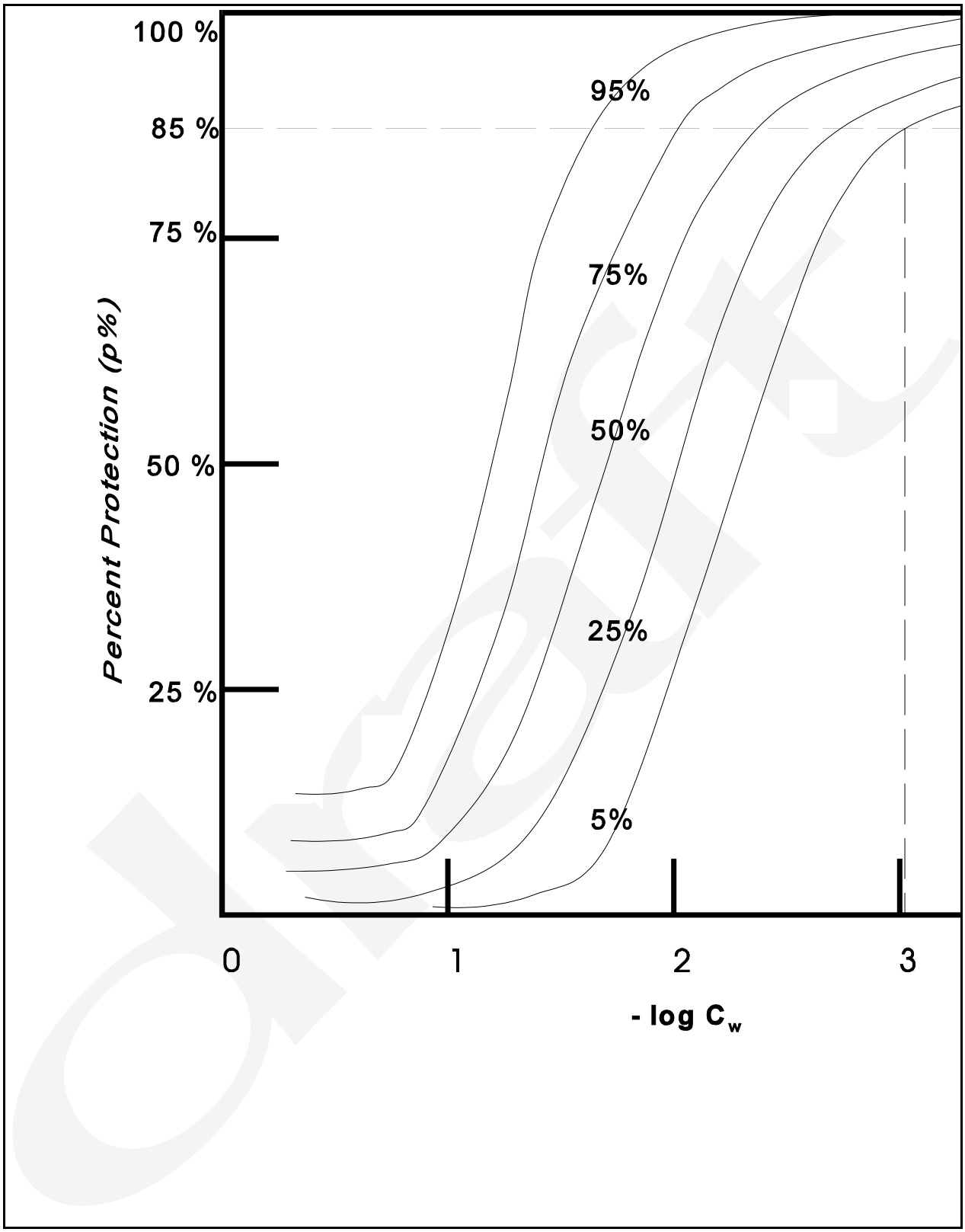

Figure 3.10 **Isopleths of the percent protection, p%, for given levels of uncertainty. A regulatory value of -log C_w=3, results in a 95% chance that the level of protection level will be at least equal to 85% (or alternatively a 5% chance that the level of protection will be less than or equal to 85%).**

4.0 SUMMARY AND DISCUSSIONS

A methodology for multimedia, multipathway and multireceptor risk assessment (3MRA) has been developed to determine regulatory constituent-specific-based exemption levels for chemicals in wastes managed in industrial Subtitle D waste management units. The concentration limits apply uniformly throughout the U.S. and are chemical specific. The methodology incorporates a two-stage Monte Carlo algorithm that allows the calculation of the uncertainty and associated confidence levels in the estimated waste concentration limits as a function of the sampling and measurement errors of input parameters, and the errors in the prediction models. The methodology has been designed: to provide an estimate of the uncertainty in the estimated measures of protection associated with a regulatory waste concentration; to provide a mechanism for accounting separately for variability and uncertainty; to provide a basis for comparing the potential benefit versus cost of future sample collection efforts; and to 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. The modeling procedure uses a forward calculation and maintains mass balance at the source. The methodology is relatively general and can be used to estimate the relative importance of controllable sources of uncertainty. As the uncertainty in the estimated measures of protection increases, it can also be used to develop regulatory schemes that can result in more conservative limits.

The assessment strategy may be implemented fully provided that all of the required components and necessary resources are available. However, the scope of its actual implementation will depend on the availability of data, computational resources, and time constraints.

5.0 REFERENCES

USEPA. Screening Survey of Industrial Subtitle D Establishments, conducted by WESTAT, Office of Solid Waste, Washington, D. C., 20460, 1987.

USEPA. Wildlife Exposure Factors Hand Book. EPA/600/p-95/002 Fa, U.S. EPA, Office of Research and Development, Washington D.C. 20480, 1993.

USEPA. Identification and Listing of Hazardous Wastes: Hazardous Waste Identification Rule (HWIR): Proposed December 21. Federal Register, pp 66344 -66469, Washington, D.C. 20460, 1995.

USEPA EPA's Composite Model for Leachate Migration with Transformation Products (EPACMTP): Background Document, Office of Solid Waste, Washington, D.C., 20460, 1996.

USEPA. Exposure Factors Hand Book. EPA/600/p-95/002 Fa, U.S. EPA, Office of Research and Development, Washington D.C. 20480, 1997.

USEPA. Guiding Principles for Monte Carlo Analysis EPA/630/R-97/001, 1997

SAB. US EPA's Science Advisory Board's Report on the Review of EPA's Composite Model for leachate Migration with Transformation Products (EPACMTP). USEPA's Science Advisory Board, Washington, D.C. 20460, 1995 .

APPENDIX A

UNCERTAINTY AND VARIABILITY IN HWIR99

The consideration of variability and uncertainty plays an important role in the HWIR99 development effort. Variability arises from the true heterogeneity of a parameter over space and/or over time. It is distinct from uncertainty which represents a lack of information or knowledge of a parameter or model either due to lack of data, or imprecise and/or insufficient measurements, or insufficient knowledge. In the case of HWIR99, variability and uncertainty of the measures of protection arises from the variability and uncertainty of the risk model input parameters, and the uncertainty of the risk model component predictions. The remainder of this appendix presents a short summary of the sources of uncertainty and variability in HWIR99, and a discussion of the importance of accounting separately for uncertainty and variability.

A.1 SOURCE OF VARIABILITY AND UNCERTAINTY

One of the principal sources of variability in the HWIR work is the variability of input parameters between sites. Example sources of variability include the between-site variability of the waste management characteristics such as area and volume, average spatial groundwater characteristics, climatic parameters, and number and type of receptors. Although spatial variability can also occur within sites, it is likely to be a significantly smaller contribution of the overall variability than the between-site variability.

There are a number of sources that contribute to the uncertainty in the prediction of the protective regulatory levels. These uncertainties can be generally classified as sampling and non-sampling errors. Sampling errors arise because the number of samples (n) where a parameter is measured (sampled) is less than the number of sites in the population (N). The magnitude of the sampling error is a function of the variability of the parameter, the sample size n, and the population size (N). In general, the magnitude of the sampling error will be proportional to the variability and inversely proportional to the sample size. Non-sampling errors are generally independent of the sample size and are generally more difficult to estimate. Examples of non-sampling errors include measurement errors, simulation model errors, errors due to non-probability samples, improper problem statements, and errors due to sampling from non-target populations.

The input parameters for the proposed framework are used to define the modeling scenario for a facility and can be grouped into four general classes: 1) variables that describe the characteristics of the waste management facility, including area and depth; 2) variables that describe the environmental conditions of the facility and its surroundings including hydrologic, hydrogeologic, meteorologic, and geochemical conditions at the site; 3) variables that describe the (physiologic and behavioral) exposure and response characteristics of the receptors; and 4) variables that describe the physical, chemical, and biochemical properties of the chemical constituents.

The first class of input parameters can exhibit variability, and uncertainty due to measurement errors and sampling errors. The second class of parameters can exhibit within and between-facility variability, and uncertainty due to data measurement errors, sampling errors, and potentially errors due to the collection of non-probability samples. The third class of parameters can exhibit between facility variability, between individual receptor variability, and uncertainty due to sampling errors, measurement errors, and potentially errors due to the collection of non-probability samples, or non-representative samples. Finally, the fourth class of parameters are characterized by variability between batches, and uncertainty due to sampling and measurement error.

There are also a number of prediction model error sources that would arise in the Monte Carlo simulation of the nationwide distributions of the protection measures, including: the mechanistic model prediction of the multimedia emission source terms from the WMU; the multimedia fate and transport modules that predict the media contaminant concentrations; the exposure models that predict the receptor dose; and the effect/response models that predict the receptor impacts. Additionally, there is the potential error of improperly stating the problem.

A.2 SEPARATING VARIABILITY AND UNCERTAINTY

Separating the effects of variability and uncertainty in estimating the nationwide probability distribution of measures of protection is important for a number of reasons. First, it permits the estimation of the uncertainty in any estimated measure of the nationwide variability of the protection measure. For example, instead of reporting the 90th percentile of the nationwide risk measure, the separation of the variability and uncertainty allows the reporting of the 95% confidence limits of the 90th percentile of the nationwide risk measure. Second, it allows the identification of sources of uncertainty that are potentially reducible so that strategies for reducing the uncertainty can be developed. Additionally, as shown in the following paragraph, it can affect the determination of whether a waste concentration meets the protection measure criteria.

The separation of uncertainty and variability can be accomplished through a two-stage Monte Carlo procedure that produces the $N_f \times N_i$ output matrix described in the previous section. How the uncertainty and variability are separated is case specific. and depends on whether the parameter is either: a) variable and certain; b) constant and uncertain; c) variable and uncertain; or d) constant and certain. To illustrate the basic elements of a two-stage Monte Carlo, and how separating variability and uncertainty can affect the regulatory limits, consider the hypothetical case where the probability distribution of the risk (R') of the nationwide receptors of concern for a given waste concentration , C_w, is lognormal so that the log of risk is normally distributed with unknown mean, μ, and known variance, σ^2:

$$R = \text{Log}(R') \sim N(\mu, \sigma^2) \qquad \textbf{(A.1)}$$

Uncertainty occurs from lack of knowledge of the true mean μ as a result of sampling error. This uncertainty is represented by a normal probability distribution with known mean $\theta = -15$ and known variance, $\tau^2 = 16$:

$$\mu \sim N(\theta, \tau^2) \qquad \textbf{(A.2)}$$

The uncertainty in the mean, as described by the probability distribution function (pdf) in equation (A.2) could have been derived in a number of ways including Bayesian (DeGroot, 1970), empirical Bayesian, or parametric bootstrap methods (Efron and Tibshirani, 1993). The variability in risk is given by $\sigma^2 = 16$, which for this example is the same as the uncertainty in risk as given by τ^2. The remainder of the discussion is based on the assumption that the protection measure is 90% of receptors protected for a target risk of 10^{-5}.

For this simple case, three cases are considered to illustrate the effects of incorporating and separating uncertainty from variability: 1) Uncertainty is included, and uncertainty and variability are separated; 2) uncertainty is included, but uncertainty and variability are not separated; and 3) uncertainty is not included.

In the first case, the separation of uncertainty and variability allows the description of the uncertainty for any given measure of the probability distribution describing the variability. In the HWIR case, the interest is in the uncertainty of the pth percentile of the nationwide risk, or more formally the upper Quth percentile

of uncertainty of the Pv^{th} percentile of variability of the log risk R. For this case, the Monte Carlo would consist of an NxM matrix of log risk realizations. Each of the M columns would be generated by first generating a value of the uncertain mean, μ, from (A.2), and then simulating N values of R from the probability distribution given by (A.1) for the given value of the uncertain mean. For each column, an estimate of the Pv^{th} percentile of variability would be estimated. The M resulting estimates of the Pv^{th} percentiles of variability for each of the M columns would then be used to estimate the uncertainty as reflected by the Qu^{th} percentile of uncertainty of the Pv^{th} percentile of variability of the log risk R

In the second case, uncertainty is not separated from variability. As a result uncertainty cannot explicitly be described for the variability. Instead the pth percentile of the nationwide risk distribution incorporates both uncertainty and variability. For this case, the Monte Carlo simulation would involve the generation of a single N*M vector of realizations, where for each N values of R correspond to a given value of the uncertain mean, μ, from (A.2). The estimate of the pth percentile of the N*M vector of realizations would incorporate both uncertainty and variability.

Finally, in the third case, uncertainty is not included in the analysis so that the distribution of nationwide risk only includes variability. For this case, the Monte Carlo simulation involves N simulations of R using (A.1), with the mean given by θ. The estimate of the pth percentile of variability from the N simulated R values would only include variability.

Figures A.1 and A.2 show the different types of results that are obtained for the three cases, depending on how uncertainty and variability are addressed.

The dashed line in Figure A.1, designated as P(u+v) corresponds to the second case. It represents the cumulative probability distribution function (cdf) of the log of risk for the given waste concentration based on a one-stage Monte Carlo. For a given risk value, the cdf provides an estimate of the percent of nationwide receptors whose risk is less than the given risk value. P(u+v) is obtained by analyzing the combined (NxM) output matrix of percent protections as a single data set, rather than by analyzing each iteration of the output matrix individually. As a result, the resulting cdf, P(u+v), incorporates both uncertainty and variability, but does not separate them. In particular, the resulting cdf shows that 96% of the receptors have risk less than 10^{-5}. On the basis of the one-stage Monte Carlo, the waste concentration would be considered protective of the specified protection measure.

The three curves labeled P(u|v)95%, P(u|v)5% and P(u|v)med in Figure A.1 correspond to the first case and illustrate the results of separating uncertainty and variability. Unlike the one-stage Monte Carlo, the two-stage Monte Carlo permits the estimation of the uncertainty in the protection measure by analyzing each iteration of the output matrix individually. Each iteration provides one estimate of the protection measure which can then be analyzed to estimate the uncertainty in the protection measure. The uncertainty can be depicted in a number of ways. In this example, the uncertainty is described by showing the 5% and 95% confidence limits for the cumulative distribution function of the log of risk. The curve that forms the lower envelope, and which is denoted by P(u|v)95%, indicates that there is a 95% chance that the actual percentage of protected receptors will be at least equal to the value indicated by the curve. Specifically, P(u|v)95% indicates that there is 95% chance that at least 80% of the receptors would have risk less than the target risk of 10^{-5}. Similar analysis can be used to show that there is an 89% chance that the measure of protection would be met for the given waste concentration; that the 90% receptor protection could be met with a 95% chance only for a risk of $10^{-3.3}$; and that the 96% receptor protection estimated by P(u+v) would be met with only a 77% chance. If the protection measure were modified by adding the additional constraint that the protection criteria would have to be met with at least a 95% confidence, then the waste concentration in the example would not qualify as protective.

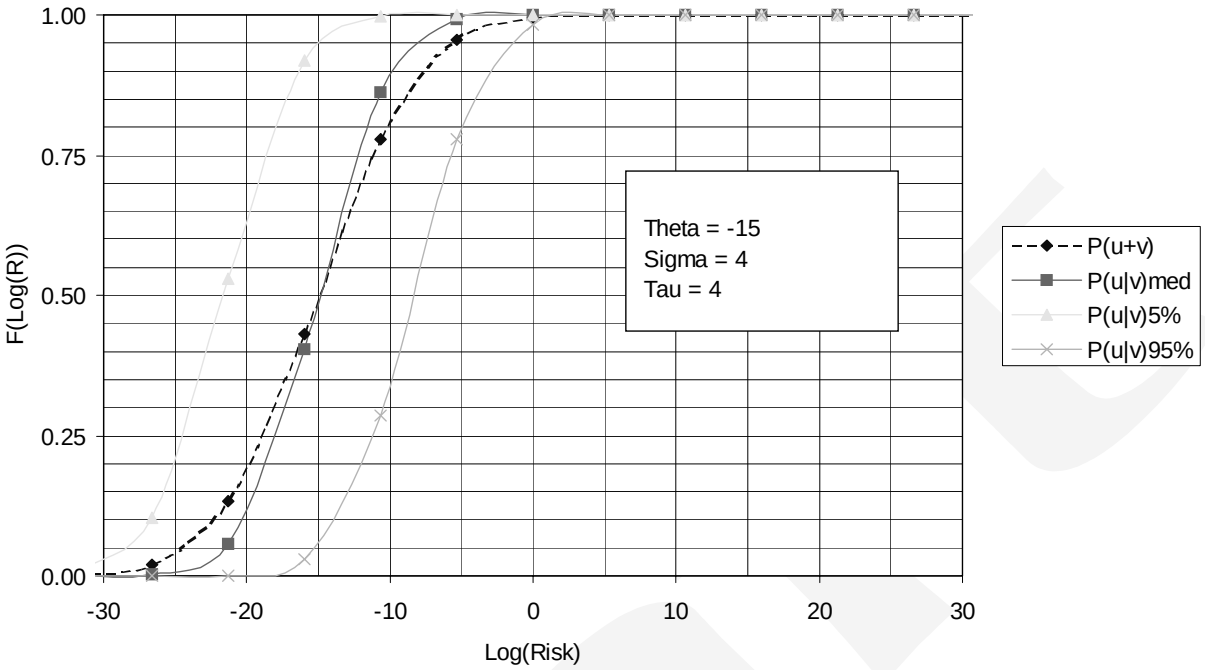

Figure A.1 Distribution of Risk under Uncertainty.

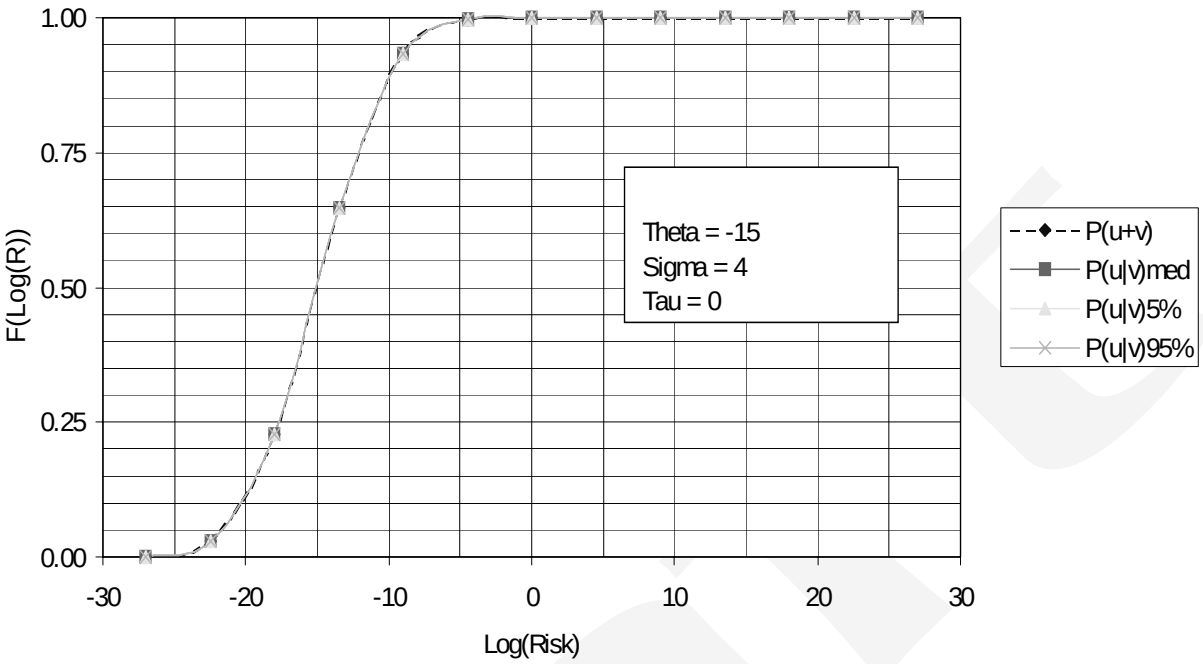

Figure A.2 Distribution of Risk under No Uncertainty.

The median curve in Figure A.1, denoted by P(u|v)med, provides an estimate of the percentage of receptors that have risk less than a specified risk if uncertainty is ignored. The same curve is shown in Figure A.2 which illustrates how the four different cdfs collapse to the median (mean) curve when the uncertainty, as represented by τ, is zero. The median (mean) curve shows that ignoring uncertainty leads to the conclusion that 99.4% of the receptors would have risk less than the target risk of 10^{-5}. Ignoring uncertainty would thus lead to accepting the waste concentration as protective.

This example illustrates the potential importance of incorporating uncertainty, and separating its effects from variability. Ignoring uncertainty and/or failing to separate uncertainty from variability prevents the characterization of the uncertainty in protection measures, and can lead to optimistic estimates of protection.

A.3 REFERENCES

DeGroot, M. H. Optimal Statistical Decisions. McGraw-Hill, Inc. New York, N.Y., 1973.

Efron, B. and R. J. Tibshirani. An Introduction to the Bootstrap. Chapman & Hill, New York, N.Y. 10003, 1993.