Variability and Uncertainty in Risk Estimation

for Brownfields Redevelopment

Arne Edward Olsen

Thesis Submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

In

Environmental Science and Engineering

Approved:

_____
Dr. Naraine Persaud, Chair

_____
Dr. Matthew Eick

_____
Dr. William Moore

_____
Sanjay V. Thirunagari

Keywords: Brownfields, Risk, Deterministic, Probabilistic

July 25, 2000
Blacksburg, Virginia

Variability and Uncertainty in Risk Estimation for Brownfields Redevelopment

Arne E. Olsen

Abstract

Various methods can be used to estimate the human health risk associated with exposure to contaminants at brownfields facilities. The deterministic method has been the standard practice, but the use of probabilistic methods is increasing. Contaminant data for non-carcinogens and carcinogens from 21 brownfields sites in Pennsylvania were collected and compiled. These were used to evaluate the performance of the deterministic and several probabilistic methods for assessing exposure and risk in relation to variability and uncertainty in the data set and input parameters. The probabilistic methods were based (a) entirely on Monte Carlo simulated input parameter distribution functions, (b) on a combination of some of these functions and fixed parameter values, or (c) on a parameter distribution function. These methods were used to generate contaminant intake doses, defined as the 90^{th}, 95^{th}, or 99.9^{th} percentile of their estimated output distribution, for the principal human exposure routes. These values were then compared with the corresponding point values estimated by the deterministic method. For all exposure routes the probabilistic intake dose estimates, taken as the 90^{th} and 95^{th} percentiles of the output distribution, were not markedly different from the deterministic values or from each other. The opposite was generally the case for the estimates at the 99.9^{th} cutoff percentile, especially for the Monte Carlo-based methods. Increasing standard deviation of the input contaminant concentration tended to produce higher intake dose estimates for all estimation methods. In pairwise comparisons with the deterministic estimates, this trend differed significantly only for the probabilistic intake doses estimated as the 99.9^{th}

percentile of the output distribution. Taken together, these results did not indicate clear and definitive advantages in using probabilistic methods over the deterministic method for exposure and risk assessment for brownfields redevelopment. They supported using the tired system for environmental risk assessment at any particular brownfields facility.

ACKNOWLEDGEMENTS

I would like to express my gratitude to Dr. Naraine Persaud for his guidance and direction to successfully complete this phase of my planned sojourn in graduate school. I would also like to thank Dr. Matthew Eick, Dr. William Moore, and Sanjay Thirunagari for the advice and support they have provided throughout this course of study. The Graduate Student Association award of funds to purchase the @Risk software package is very much appreciated. Lastly, I would like to thank my parents, Arne and Becky Olsen, for their support and encouragement during my graduate studies.

TABLE OF CONTENTS

Abstract i

ACKNOWLEDGEMENTS iii

TABLE OF CONTENTS iv

LIST OF TABLES v

LIST OF FIGURES viii

INTRODUCTION 1

LITERATURE REVIEW 6

2.1 Brownfields 6
2.2 Risk Assessment Methodology 9
2.2.1 Basic Concepts 9
2.2.2 Variability and Uncertainty in Risk Estimation 18
2.2.3 Deterministic Risk Assessment 20
2.2.4 Probabilistic Risk Assessment 23

MATERIALS AND METHODS 33

3.1 Brownfields Sites and Contaminant Data 33
3.2 Deterministic Risk Estimates 41
3.3 Probabilistic Risk Estimates 43
3.4 Statistical Analysis 46

RESULTS AND DISCUSSION 50

SUMMARY AND CONCLUSIONS 75

References 79

APPENDIX

Contaminant concentrations, their descriptive statistics and Land's UCL_{95}, calculated mass transfer parameters, and @Risk best fits for soil, air, and groundwater concentration distributions. 85

Vita 94

LIST OF TABLES

Table 1. Twenty-one sites taken from Pennsylvania's Land Recycling Program list of completed brownfields.....34

Table 2. Parameter values for the intergerated soil to air emission and dispersion models.....39

Table 3 . Parameter values for estimating groundwater concentraiton from soil concentration.....40

Table 4. Values of exposure factors for the deterministic and Monte Carlo/deterministic analyses.....42

Table 5. Recommended density distributions and @ Risk 4.0 function parameters (in parenthesis) for each exposure variable in the Monte Carlo and Monte Carlo/deterministic analysis.....44

Table 6. Mean and standard deviation of exposure variables used for the two-point-based risk estimation method.....45

Table 7. Contaminant/site combinations and exposure routes for which intake dose for non-carcinogens were estimated.....47

Table 8. Contaminant/site combinations and exposure routes for which intake dose for carcinogens were estimated.....47

Table 9. Methods of risk estimation and their abbreviations. For each of the probabilistic methods the 90^{th}, 95^{th} or 99.9^{th} percentile value of the estimated intake dose distribution was used for comparison.....51

Table 10. Intercept, slope and coefficient of determination (r^2) for modified linear regression of intake doses by soil ingestion estimated by different methods, versus the standard deviation of the measured soil concentration of non-carcinogens 69

Table 11. Intercept, slope and coefficient of determination (r^2) for modified linear regression of intake doses by dermal absorption estimated by different methods, versus the standard deviation of the measured soil concentration levels of non-carcinogens 69

Table 12. Intercept, slope and coefficient of determination (r^2) for modified linear regression of intake doses by inhalation estimated by different methods, versus the standard deviation of the modeled air concentration of non-carcinogens 70

Table 13. Intercept, slope and coefficient of determination (r^2) for modified regression of intake doses by groundwater ingestion estimated by different methods, versus the standard deviation of the modeled groundwater concentration of non-carcinogens 70

Table 14. Intercept, slope and coefficient of determination (r^2) for modified linear regression of intake doses by soil ingestion estimated by different methods, versus the standard deviation of the measured soil concentration levels of carcinogens 72

Table 15. Intercept, slope and coefficient of determination (r^2) for modified linear regression of intake doses by dermal absorption estimated by different methods, versus the standard deviation of the measured soil concentration of carcinogens. 73

Table 16. Intercept, slope and coefficient of determination (r^2) for regression of intake doses by groundwater ingestion estimated by different methods, versus the standard deviation of the modeled groundwater concentration of carcinogens. 74

LIST OF FIGURES

Figure 1. Mean estimated intake dose for the soil ingestion exposure route for non-carcinogens. Method abbreviations are defined in Table 9. Values not annotated with the same letter are significantly different at $p \leq 0.05$. Bars represent the standard error 53

Figure 2. Mean estimated intake dose for the dermal exposure route for non-carcinogens. Method abbreviations are defined in Table 9. Values not annotated with the same letter are significantly different at $p \leq 0.05$. Bars represent the standard error 55

Figure 3. Mean estimated intake dose for the inhaletion exposure route for non-carcinogens. Method abbreviations are defined in Table 9. Values not annotated with the same letter are significantly different at $p \leq 0.05$. Bars represent the standard error 56

Figure 4. Mean estimated intake dose for the groundwater ingestion exposure route for non-carcinogens. Method abbreviations are defined in Table 9. Values not annotated with the same letter are significantly different at $p \leq 0.05$. Bars represent the standard error 58

Figure 5. Mean estimated intake dose for the soil ingestion exposure route for carcinogens. Method abbreviations are defined in Table 9. Values not annotated with the same letter are significantly different at $p \leq 0.05$. Bars represent the standard error 62

Figure 6. Mean estimated intake dose for the dermal exposure route carcinogens. Method abbreviations are defined in Table 9. Values not annotated with the same letter are significantly different at $p \leq 0.05$. Bars represent the standard error 63

Figure 7. Mean estimated intake doses for the groundwater ingestion exposure route for carcinogens. Method abbreviations are defined in Table 9. Values not annotated with the same letter are significantly different at $p \leq 0.05$. Bars represent the standard error 64

CHAPTER 1

INTRODUCTION

The federal brownfields program was initiated in 1995 to stimulate redevelopment of the economic base of communities that have been negatively impacted by the loss of jobs and tax base that were previously associated with brownfields facilities (Davis and Margolis, 1997). Then Vice-President Al Gore emphasized this in June 1996 when he wrote in reference to brownfields (Davis and Margolis, 1997).

> *" . . . As we embark on this exciting new age, we cannot forget the many people and places left behind; people and places that want to join in the progress we are making. The estimated hundreds of thousands of abandoned and contaminated properties that are littered across our poorest communities – known as 'brownfields' – present a significant barrier to economic revitalization in our nation's cities. By encouraging the cleanup and redevelopment of these brownfields, the Clinton Administration is forging new ways to empower distressed communities and create jobs for their residents . . . "*

Once brownfields have been redeveloped and placed back into the economy, localities can collect real-estate taxes and/or income taxes. The added revenues allow the communities to improve infrastructure, which provide new economic development opportunities thereby reversing the negative effects of urban decay. At the same time, brownfields redevelopment slows the development of virgin, undeveloped properties (greenfields), thus reducing the blight of urban sprawl.

The environmental contamination associated with a given brownfields facility creates a liability for potential redevelopers. This liability can eliminate the possibility of obtaining loans for the redevelopment of the facility. The financial risk of investing in brownfields is therefore a direct function of the potential for deleterious environmental effects posed by the contamination at the site to the surrounding community (Davis and

Margolis, 1997). This potential for harm is quantified by assessing the human health and environmental risk posed by the contamination at the site. Usually, as the health and environmental risks increase, the costs of removing these risks escalate and the likelihood of investment in these properties is decreased (Davis and Margolis, 1997).

The human health and environmental risks posed by contaminants at a brownfields site are calculated using risk assessment algorithms that relate the level of contamination, toxicity, and exposure at the site (Davis and Margolis, 1997). The primary objective of the risk assessment process is to determine whether current levels of contamination at the facility can result in unacceptable deleterious health effects and consequently require cleanup action or mitigation before redevelopment can proceed. Currently, the risk assessment process prescribed by the United States Environmental Protection Agency (USEPA) uses deterministic values for levels of contamination and exposure factors (i.e., intake rates, exposure time, exposure duration, body weight, etc.). These levels of contamination and exposure factors are then used to determine the exposure level or intake dose. The intake dose is then combined with established toxicity values to determine the potential for future deleterious health effects (USEPA, 1989).

Variability and uncertainty are always associated with risk assessment since they are intrinsic to the contaminant levels, toxicity values, and the exposure factors used in the calculations. A quantitative assessment of the uncertainty and variability is an integral part of risk assessment (USEPA, 1989). In order to take into account such uncertainty and variability USEPA designed its deterministic risk assessment methodology to quantify the risk associated with the reasonable maximum exposure (RME). The RME is a conservative exposure level chosen within the 90^{th} to 99.9^{th}

percentile of the estimated range of possible exposures (USEPA, 1989). It has been suggested that the assumptions incorporated in the USEPA risk assessment process are too conservative and result in overestimating the RME and the estimated risk to the maximally exposed individual (Bogen, 1994). Additionally, combining several such upper-bound estimates in a multiplicative risk model would overestimate risk through the process termed as compounded conservatism (or "safety creep"). The outcome of the model would represent an upper bound that is greater than the input upper bounds (Bogen, 1994). Therefore, the risk assessment process may produce unreasonably high estimates of the risk posed by contaminants at the facility. If this occurs when assessing risk for brownfields, their redevelopment may be artificially stymied.

Recently introduced, probabilistic-based risk assessments are increasingly being used at Superfund sites to provide more complete estimates for the RME and its associated risk. More complete estimates for the RME can lower remediation costs while still ensuring protection of human health and the environment (USEPA, 1999a). Unlike the deterministic approach to risk assessment that only provides an upper-bound estimate of risk, probabilistic-based risk assessments are feasible and cost-effective and can generate a more complete estimate of risk by better quantifying the degree of variability and uncertainty. However, Smith (1994) found that the deterministic approach and the probabilistic approach gave consistent risk estimates. Additionally, he found that the RME values for the deterministic approach were near the 95^{th} percentile of the probabilistic estimates for the exposure routes he studied.

The probabilistic approach uses values of contaminant levels and exposure factors from probability distribution functions (pdfs) instead of the deterministic point estimates

in the traditional risk assessment approach. This method eliminates the problem of compounded conservatism associated with the deterministic method. The pdfs are generated using Monte Carlo simulations. Once the appropriate pdfs have been developed, values are selected at random from these distributions. These are then incorporated into the risk algorithms to calculate a single value of the environmental risk (Dakins et al., 1996). This process is then repeated thousands of times to develop a cumulative distribution of environmental risk. A value in the 90^{th} to 99.9^{th} percentile of this distribution is then used as the final risk value (USEPA, 1999a).

The Monte Carlo-based risk assessment method allows for sensitivity analysis to determine which parameters have the least impact on the risk outcome of a simulation. Pearson's r or Spearman's ρ can be used to test for sensitivity of the risk estimates to each parameter (Hamby, 1994; Palisade, 2000). The non-significant parameters can then be treated as deterministic variables thus reducing the number of distributions that must be developed and the number of Monte Carlo simulations that must be performed.

A less computationally tedious probabilistic method for estimating risk is available but has never been applied in environmental risk assessment. This is the "two-point method" developed by Rosenblueth (1981), and Guymon et al. (1981). This method is based on estimates for the moments of a probability density distribution (Rosenblueth, 1975). The first two moments of the probability density distributions of the input random variables in the risk algorithms are used to estimate the mean and variance of the risk outcome. The two-point method requires less iterations than the Monte Carlo-based approach, thereby reducing the calculation-checking problem. It also reduces the problem of compounded conservatism. Use of Monte Carlo generated

distributions or the two-point method can provide better and more meaningful quantification of the variability and uncertainty associated with the risk estimates than does the deterministic approach (USEPA, 1989).

The applicability and advantage of the probabilistic approach namely, the Monte Carlo-based and the two-point method, in quantifying variability and uncertainty in brownfields risk assessments need to be investigated. Since brownfields sites generally have fewer contaminants at lower concentrations than Superfund sites, there is the possibility that the Monte Carlo-based and the two-point method may not produce significantly different estimates for the reasonable maximum exposure compared to the deterministic method. However, if it can be shown that Monte Carlo-based and/or the two-point based risk assessments can better capture the uncertainty and variability inherent to the risk assessment process, the costs of remediation at brownfields sites can be reduced, and make brownfields redevelopment more attractive.

The primary objective of this study was to compare exposure and risk estimates obtained using the deterministic, Monte Carlo-based, and the two-point risk analysis method in relation to the variability and uncertainty in contaminant levels at selected brownfields facilities. Additionally, the sensitivity of the Monte Carlo risk estimates to the input parameters and their distributions were analyzed. Non-significant parameters identified by these analyses were treated as deterministic and used to develop and evaluate a hybrid Monte Carlo/deterministic method.

CHAPTER 2

LITERATURE REVIEW

2.1 Brownfields

The United States Environmental Protection Agency (USEPA, 1995) defines brownfields as: "abandoned, idled or underused industrial and commercial sites where expansion of redevelopment is complicated by real or perceived environmental contamination that can add cost, time, or uncertainty to a redevelopment project." Davis and Margolis (1997) categorized brownfields as:

1) Sites that, despite needed remediation, remain economically viable, due to sufficient market demand.
2) Sites that have some development potential, provided financial assistance or other incentives are available.
3) Sites that have extremely limited market potential even after remediation.
4) Currently operating sites that are in danger of becoming brownfields because historical contamination will ultimately discourage new investment and lending.

This classification system indicates that brownfields redevelopment is a voluntary market-driven process. Because there is substantial risk associated with investing in a brownfields site, redevelopment will only be undertaken when there is a potential for substantial return on the investment (Davis and Margolis, 1997). Therefore, only the first two categories of brownfields are economically viable. Viable brownfields are considered as under-utilized properties with actual or perceived environmental liabilities that, due to their inherently positive market attributes, may be economically redeveloped into productive assets (Davis and Margolis, 1997). Therefore, until such time as the market causes them to have value greater than the redevelopment cost, many brownfields will remain non-viable (Davis and Margolis, 1997).

The United States Government General Accounting Office estimates that there are 130,000 to 450,000 potential brownfields in the United States (Davis and Margolis, 1997). The cost to investigate and remediate all of these brownfields may be as high as $650 billion (Davis and Margolis, 1997). Although this estimated cost is high, the loss of tax dollars and lost wages is also high. The United States Conference of Mayors conservatively estimated that their annual loss of tax revenues for 33 cities due to brownfields sites totaled $121 million. Using more realistic estimates, the projected losses were closer to $386 million (Davis and Margolis, 1997).

The stigma associated with environmental contamination at industrial and commercial brownfields reduces their market value and likelihood for redevelopment. This perception indirectly reduces employment and the tax base of the community in which they are located. In this context, Davis and Margolis (1997) considers stigma to be a persistent prejudice that any environmental contamination makes a property flawed, blemished, discredited, or spoiled. This prejudice exists primarily because of the real possibility that a site is contaminated with a chemical or material that has the potential to negatively impact human health. The effects of this stigma on a community with brownfields are: (a) reduced potential for redevelopment because of fear of liability, resulting in loss of potential property-tax revenues, (b) unemployment due to reduced job base, (c) reduced need for public services and infrastructure, and (d) the disincentive for attracting new development (Davis and Margolis, 1997).

Available remediation standards can be used to determine when a brownfields site has been adequately remediated and thus remove the environmental stigma associated with the site (Davis and Margolis, 1997). Both remediation to background, or to non-

detectable levels of contamination, are often protective of human health and the environment; however, this approach often removes more contamination than is necessary resulting in increased remediation costs and decreased likelihood of redevelopment (Davis and Margolis, 1997).

To address the brownfields issues, the United States Environmental Protection Agency (USEPA) issued its "Brownfields Action Agenda" in January 1995. The action agenda was composed of four parts. First, the USEPA would put in place 50 pilot projects. Each project would receive $200,000 dollars over two years to help fund site evaluation and remediation (USEPA, 1995). Second, USEPA would clarify the liability and cleanup issues for prospective purchasers, lenders, and property owners. Included in this phase was the archiving of 24,000 sites that were slated for no further action from the CERCLIS database, (CERCLIS is the Comprehensive Environmental Response, Compensation, and Liability Information System) which is the repository for information on hazardous waste sites, site inspections, preliminary assessments, and remediated hazardous waste (USEPA, 1995). Additionally, the USEPA would issue guidance that required the consideration of future land use when selecting remediation technologies at sites on the National Priorities List. Third, the action agenda would encourage public participation and community involvement. Fourth, the USEPA would establish an environmental education program to develop an environmentally conscious workforce (USEPA, 1995). Since the release of the “Brownfields Action Agenda” in 1995 all of the milestones have been accomplished including clarification of liability and cleanup issues, partnerships and outreach, job training, and brownfields pilots (USEPA, 2000a).

Five communities in Virginia received pilot project funds to aid in redevelopment of brownfields. These were Petersburg, Newport News, Shenandoah, Cape Charles, and Richmond. Petersburg and Newport News planned to use the funds to redevelop depressed industrial areas in the inner city to spur economic development (USEPA, 2000b; USEPA, 1999b). Shenandoah planned to use the funds to redevelop a former iron furnace into recreational center and historical park (USEPA, 1998). Cape Charles used the funds to help redevelop a 25-acre town dump and surrounding lands into a Sustainable Technology Industrial Park, a coastal dune preserve, and other natural areas. The project helped create 82 new jobs in a community where 27 percent of the 13,000 residents live below the poverty level (USEPA, 1999c). Richmond used the funds to assess and provide plans for the remediation of a 4.5-acre parcel of land. Whithall-Robins, a pharmaceutical company, subsequently spent nearly $2 million dollars to redevelop the land and was then given the site. The remediation allowed further development of the land that resulted in the creation of 250 new jobs and allowed 100 positions to remain in the Richmond area. The new facility generates an average of $100,000 per year in tax revenues (USEPA, 1999d).

2.2 Risk Assessment Methodology

2.2.1 Basic Concepts

The Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) of 1980 established the federal program that guides response to uncontrolled releases of hazardous substances into the environment. As authorized by CERCLA, the USEPA established protocols for assessing and remediating sites contaminated with

hazardous substances (USEPA, 1989). An essential part of this protocol was risk assessment. Risk assessment can be used to set remediation limits, lower remediation costs, lessen the stigma, and increase chances for redevelopment of a contaminated facility. Risk assessment is intended to be a process of gathering, organizing, and analyzing data to establish the probability of adverse health impacts from hazardous site conditions (Williams et al., 2000). The USEPA based its guidance for risk assessment on the National Research Council's *Risk Assessment in the Federal Government: Managing the Process* (USEPA, 1989), already established guidance from other federal agencies such as OSHA (Occupational Safety and Health Administration), FDA (Food and Drug Administration), and CPSC (Consumer Product Safety Commission), and its own accumulated experience conducting risk assessments at hazardous waste sites. The final guidance issued by USEPA was designed to (a) provide an estimate of risk and determine whether further action was warranted, (b) provide estimates of the highest level of a contaminant at a site that is not harmful to human health and the environment, (c) provide a means for comparing different possible remedial options, and (d) provide consistency in the methodology for determining human health risk from environmental contaminants (USEPA, 1989).

The National Research Council's publication entitled *Risk Assessment in the Federal Government: Managing the Process* (NRC, 1983) established the process currently used for conducting risk assessments and defined many of the operational terms. This document, sometimes referred to as the "Red Book", also established the basic risk assessment paradigm (NRC, 1994). This paradigm comprises a four-step process, namely, hazard identification, dose-response assessment, exposure assessment,

and risk characterization (Williams et al., 2000). The first two are sometimes grouped as toxicity assessment.

Hazard identification involves the determination of the potential impacts that exposure to a contaminant may have on the health of the exposed individual. These health impacts may be, but are not limited to, carcinogenesis, mutagenesis, neurotoxicity, hematotoxicity, nephrotoxicity, dermal and ocular toxicity, pulmonotoxicity, immunotoxicity or hepatotoxicity (Williams et al., 2000). Although it is known that exposure to contaminants under some conditions can be toxic to humans, there is little direct evidence of causal relationships between exposure and toxicity for most chemicals. The most accurate method of determining toxicity would be to test the effects of exposure to contaminants directly on human subjects, which is not possible for obvious reasons. Therefore most assessments of health hazards are performed indirectly using epidemiological studies, studies of accidental exposures, studies on laboratory animals, and comparison with pharmacokinetics of similar chemicals with known hazards (Williams et al., 2000).

The next step in the risk assessment paradigm is the dose-response assessment or the relationship between exposure level and the incidence and severity of deleterious health effects. Dose-response assessment not only takes into account the relationship of exposure to toxicity, but also the effects of exposure patterns, age, sex, and lifestyle factors (NRC, 1994). The dose-response relationship is often arrived at by extrapolation from high to low doses and from test animals to humans subjects (Williams et al., 2000).

Exposure assessment is the third step of the risk assessment paradigm and quantifies the intensity, frequency, and duration of human exposure to environmental

contaminants. Quantifying exposure can be simple or extremely complex. The concentration of a contaminant, to which an individual is exposed, can be directly measured by collecting field samples and determining the concentration in the samples in an analytical laboratory. Exposure concentrations can also be modeled indirectly using the known physiocochemical relationships that control the movement of contaminants from source to receptor (USEPA, 1989).

The exposure concentration is dependent on the exposure pathway or any combination of potential exposure pathways (NRC, 1994). An exposure pathway describes how a given contaminant reaches and enters the human body. The exposure pathway therefore links the source and the type of contamination to the potentially exposed individual or population. For an exposure pathway to be complete there must be a source of contamination, a means of transport to the receptor, an exposure route into the receptor, and a potential for exposure by the given route (USEPA, 1989). Exposure pathways can be defined for contaminant movement through any combination of the air, water, and soil. Principal human exposure routes are ingestion, absorption and inhalation.

The exposure frequency and duration, which depends on the exposure scenario, complicates the exposure assessment. Whether the exposed individual resides, works, or performs any other activity at the exposure point, dictates the estimated exposure frequency and duration (USEPA, 1989). Generally, the residential scenario is associated with the greatest exposure and is therefore considered the most conservative exposure scenario. The residential exposure is therefore often considered for CERCLA sites.

However, in situations where the possibilities of future residential activities at a site are exceedingly small, the residential scenario will overestimate potential exposure or risk.

The final step of the risk assessment paradigm is risk characterization. In this step, the information from the first three steps is brought together to quantify the probability for deleterious health effects. Included in this step is the communication of the potential risk. This communication should include quantitative statements of all the variability and uncertainty that are associated with each of the three preceding steps and the final risk estimate (USEPA, 1989).

The methodologies for conducting risk assessments are described in the operations manual entitled "Risk Assessment Guidance for Superfund Volume I: Human Health Evaluation Manual Part A" (USEPA, 1989). After appropriate data collection, the risk is calculated for each individual contaminant. The risk for both carcinogenic effects and noncarcinogenic effects of a contaminant can be calculated. The equation for calculating carcinogenic risk is the product of the carcinogenic slope factor for the contaminant and its chronic daily intake dose expressed as the mass of a contaminant averaged over 70 years per unit body weight of the receptor. The slope factor converts the chronic daily intake value directly into the increase in the individual risk of developing cancer (USEPA, 1989). The non-carcinogenic risk equation is the quotient of the intake of a contaminant of concern averaged over a specified period (also termed as the intake dose and expressed as mass per unit body weight per day) to its toxicity threshold value termed as the reference dose (RfD). The RfD is the intake of a contaminant below which there is not a substantial likelihood of toxicological effects

(USEPA, 1989). The ratio is called the hazard quotient or hazard index. Risk or hazard is therefore determined by the following equations (USEPA, 1989):

Carcinogen: $\text{Individual Lifetime Excess Cancer Risk} = \text{Intake} \times \text{Slope Factor}$

Non-carcinogen: $\text{Hazard} = \frac{\text{Intake}}{\text{ReferenceDose}}$

In assessing risk for multiple contaminants the risk value is taken as the sum of the individual risk values (USEPA, 1989).

For carcinogens, the slope factor represents an upper bound estimate of the probability of a carcinogenic response per unit intake dose of a contaminant over an average lifetime of 70 years. The slope factor multiplied by the chronic daily intake value gives a maximum probability that an exposed individual will develop cancer from exposure to a contaminant over his or her lifetime (USEPA, 1989). Slope factor values are established by the Carcinogen Risk Assessment Verification Endeavor (CRAVE) workgroup, which was established to resolve the conflicting values that were being used by different program groups within USEPA (USEPA, 1989). These values are archived in USEPA's Integrated Risk Information System (IRIS) database.

For non-carcinogens the RfD is determined using the No Observed Adverse Effect Level (NOAEL) or Lowest Observed Adverse Effect Lever (LOAEL) threshold values taken from toxicological studies of dose-response in test animals. The NOAEL is the highest level of intake by a given route that, statistically, has no adverse health impact on the treated population in relation to a control population in the study. The LOAEL is the lowest level of exposure that, statistically, has an adverse health impact on the subject population in relation to a control population (Williams et al., 2000). The NOAEL is then divided by uncertainty factors that reduce the threshold value by several orders of

magnitude. These uncertainty factors account for differences within the exposed population, differences between test animals and the exposed human population, values that were derived from subchronic studies versus chronic studies, and NOAEL values that were determined by extrapolation from the LOAEL values. An additional factor of 10 may be applied based on professional judgement to account for unspecified uncertainties (USEPA, 1989).

Both carcinogenic slope factors and non-carcinogenic RfD's are available in the IRIS database that is accessible on the USEPA Internet site http://www.epa.gov/iris/. Such values are also published in the Health Effects Assessment Summary Tables (HEAST) and the Agency for Toxic Substances and Disease Registry (ATSDR) toxicological profiles. Values from the IRIS database are preferred as they have undergone greater scrutiny than the other values (USEPA, 1989).

Both carcinogenic and non-carcinogenic risk calculations incorporate a value for intake, also termed as the administered dose or intake dose of the chemical. The generic equation for calculating the chemical intake dose by a given route is (USEPA, 1989):

$$I = \frac{C \times CR \times EFD}{BW} \times \frac{1}{AT}$$

Where: I = intake dose (mg kg^{-1} day^{-1}); C = concentration in the medium of interest (mg $liter^{-1}$, mg kg^{-1}, mg m^{-3}) ; CR= contact rate (liter day^{-1}, kg day^{-1}, m^3 day^{-1}); EFD = exposure frequency and duration and comprises two terms: exposure frequency (day $year^{-1}$) and exposure duration (years); BW = body weight (kg); and AT = averaging time (day) (USEPA, 1989). The AT may be the same as the exposure duration or taken as some average lifetime. Because of the high uncertainty associated with the measurements of concentration, USEPA (1989) requires the use of the 95^{th} percentile of

the estimated range of the mean values for the chemical concentration in the generic equation (USEPA, 1989). Contact rate is an estimate of the amount of contaminated medium contacted per unit time. Exposure frequency and duration are combined to determine the total time of exposure. The USEPA suggests that these values be determined from site specific information. However, default assumptions that are based on the exposure scenario (residential or industrial) are often used to specify these values. Body weight is estimated using the average body weight of the exposed population (USEPA, 1989). The averaging time for non-carcinogens is the same as the product of exposure frequency and duration in days. For carcinogens it is taken as an average lifetime of 70 years expressed in days.

For each exposure route there is a unique version of the generic equation that relates the amount of contamination in the media of concern to the intake dose expressed as mass of chemical per unit body weight per day. Four common intake dose models for residential exposure are listed below (USEPA, 1989):

(1) Residential Exposure: Ingestion of Chemicals in Soils

$$\text{Intake}(\text{mg kg}^{-1}\ \text{day}^{-1}) = \frac{\text{CS} \times \text{IR} \times \text{CF} \times \text{FI} \times \text{EF} \times \text{ED}}{\text{BW} \times \text{AT}}$$

CS	=	Concentration in soil (mg kg^{-1})
IR	=	Ingestion rate (mg soil day^{-1})
CF	=	Conversion factor (10^{-6} kg mg^{-1})
FI	=	Fraction of IR attributable to contaminated soil (unitless)
EF	=	Exposure frequency (days $year^{-1}$)
ED	=	Exposure duration (years)
BW	=	Body weight (kg)
AT	=	Averaging time ($\text{ED} \times 365\ \text{day year}^{-1}$ for non-carcinogens and $70\ \text{years} \times 365\ \text{days year}^{-1}$ for carcinogens)

(2) Residential Exposure: Dermal Contact with Chemicals in Soil

$$\text{Intake}(\text{mg kg}^{-1}\ \text{day}^{-1}) = \frac{\text{CS} \times \text{CF} \times \text{SA} \times \text{AF} \times \text{ABS} \times \text{EF} \times \text{ED}}{\text{BW} \times \text{AT}}$$

CS = Concentration in soil (mg kg^{-1})
CF = Conversion factor (10^{-6} kg mg^{-1})
SA = Skin Surface area available for contact (cm^2)
AF = Soil to skin adherence factor (mg soil cm^{-2})
ABS = Absorption factor (unitless)
EF = Exposure frequency (days $year^{-1}$)
ED = Exposure duration (years)
BW = Body weight (kg)
AT = Averaging time ($\text{ED} \times 365\,\text{day year}^{-1}$ for non-carcinogens and $70\ \text{years} \times 365\ \text{days year}^{-1}$ for carcinogens)

(3) Residential Exposure: Inhalation of Airborne Chemicals

$$\text{Intake}(\text{mg kg}^{-1}\ \text{day}^{-1}) = \frac{\text{CA} \times \text{IR} \times \text{ET} \times \text{EF} \times \text{ED}}{\text{BW} \times \text{AT}}$$

CA = Concentration in air (mg m^{-3})
IR = Inhalation rate (m^3 $hour^{-1}$)
ET = Exposure time (hours day^{-1})
EF = Exposure frequency (days $year^{-1}$)
ED = Exposure duration (years)
BW = Body weight (kg)
AT = Averaging time ($\text{ED} \times 365\,\text{day year}^{-1}$ for non-carcinogens and $70\ \text{years} \times 365\ \text{days year}^{-1}$ for carcinogens)

(4) Residential Exposure: Ingestion of Chemicals in Drinking Water

$$\text{Intake}(\text{mg kg}^{-1}\ \text{day}^{-1}) = \frac{\text{CW} \times \text{IR} \times \text{EF} \times \text{ED}}{\text{BW} \times \text{AT}}$$

CW = Concentration in water (mg $liter^{-1}$)
IR = Ingestion rate (liter day^{-1})
EF = Exposure frequency (days $year^{-1}$)
ED = Exposure duration (years)
BW = Body weight (kg)
AT = Averaging time ($\text{ED} \times 365\,\text{day year}^{-1}$ for non-carcinogens and $70\ \text{years} \times 365\ \text{days year}^{-1}$ for carcinogens)

2.2.2 Variability and Uncertainty in Risk Estimation

Variability and uncertainty are intrinsic to each input parameter of the intake dose algorithms. The USEPA has provided definitions for variability and uncertainty as they apply to risk assessment. These definitions have been adopted by most of the risk assessment community. They are (Finley and Paustenbach, 1994; NRC, 1994; Dakins et al., 1996; USEPA, 1999):

Variability: the true heterogeneity or diversity in a population.

Uncertainty: a lack of knowledge regarding a population.

Uncertainty leads to inaccurate or biased estimates whereas variability leads to imprecise estimates. Uncertainty can be reduced with additional data. However, variability can only be better understood by the collection of additional data (Byrd and Cothern, 2000).

Sources of variability can be classified into three categories: spatial variability, temporal variability and inter-individual variability. The National Research Council (NRC, 1994) lists four ways to deal with these types of variability in risk estimation: (1) ignore the variability, (2) disaggregate the variability, (3) use the average, and (4) use maximum or minimum values. The first method is used when the variability is relatively low. The use of mathematical models or separating data into subpopulations are appropriate ways to disaggregate the variability in order to better understand or reduce it. Using the sample average value is similar to ignoring variability except when the sample average value is a good estimator of the true population mean. The fourth approach is the most common method of dealing with variability in risk assessment (NRC, 1994). However, the USEPA (1992a) cautions that the use of many extreme values could result in an inflated risk estimate.

Uncertainty is divided into three categories by the USEPA(1992a): (1) uncertainty regarding missing or incomplete information needed to fully define the exposure level and dose, (2) uncertainty regarding some input parameter, and (3) uncertainty regarding gaps in scientific theory required to make predictions on the basis of causal inferences. In the context of this study, uncertainty refers primarily to parameter uncertainty. It is assumed that the exposure route is complete and known and that the exposure model is correct as defined by USEPA. Parameter uncertainty can be further categorized into random errors, systematic errors, and use of generic or surrogate data (Morgan and Henrion, 1990).

All estimates of parameter values contain both variability and uncertainty that confound one another (USEPA, 1997). If a quantity having high variability is assumed to be invariant, then even the most exact measurement scheme will appear to generate very uncertain results. In this case, the faulty measurement is ascribed to uncertainty when in fact it is due to variability (USEPA, 1997). For example, in estimating the average daily dose from ingestion of contaminated drinking water, it is possible to measure an individual's daily water consumption exactly, thereby eliminating uncertainty in the measured daily dose. The daily dose still has inherent day-to-day variability, due to changes in the individual's daily water intake or the contaminant concentration in water. However, it is impractical to measure the individual's dose every day. It has to be estimated based on sample measurements. Because the individual's true average is unknown, it is uncertain how close the estimate is to the true value. Thus, the variability across daily doses has been translated into uncertainty (USEPA, 1997).

Taking conservative point estimates of the parameters and variables associated with exposure and toxicity provides a consistent, straightforward, and practical approach to dealing with variability and uncertainty in risk estimation. However, the USEPA mandate to be protective of human health and the environment, including sensitive sub-populations, contribute to a high degree of caution and conservatism in the estimation of risk. This approach has been termed the "better safe than sorry" response to variability and uncertainty (Byrd and Cothern, 2000). To be protective of a large portion of the population, USEPA chooses to bias on the side of safety.

2.2.3 Deterministic Risk Assessment

In general, deterministic models of physical processes consist of a function or a set of functions of one or more variables. In the deterministic approach to risk assessment the function variables are taken as the mean, or some upper bound of a population distribution, or the distribution of sample means. The deterministic approach therefore ignores the actual variability and uncertainty that are associated with the individual input variables (Guymon and Yen, 1990). Variability and uncertainty are lumped by using conservatively chosen values for the density distribution of the input variables such that the estimated risk represents some upper bound value. However the range of possible risk values cannot be specified and therefore, no percentile ranking can be assigned to this upper bound value (USEPA, 1999).

The most important step in the deterministic risk assessment is estimation of the intake dose or reasonable maximum exposure (RME). USEPA defines the RME as the highest exposure that is reasonably expected to occur for current and future use at a site

(USEPA, 1989). The RME is generally estimated for each exposure route. However, multiple exposure routes can be combined in calculating the RME (USEPA, 1989).

Values for the inputs needed to estimate the RME and the final risk value by the deterministic method can be found in the "Exposure Factors Handbook" (USEPA, 1997) or in other guidance documents that have been released by the USEPA or State Environmental Agencies. The concentration term is determined from site specific sampling data. The mean and standard deviation of these data are used to determine an upper (usually the 95th percentile) fiducial (confidence) limit for the sample mean. This practice is intended to ensure that the RME is not underestimated (USEPA, 1989).

The process to determine this upper fiducial limit for the mean has been laid out in *Supplemental Guidance to RAGS: Calculating the Concentration Term* released by USEPA in May 1992 (USEPA, 1992b). This document outlines the method to estimate an upper bound of log transformed data. The method provided is taken from Land (1971) who derived the uniformly most accurate unbiased confidence interval procedure.

To determine the 95th percentile fiducial limit for the mean USEPA suggests that the contaminant concentration data should be tested to determine whether the data are normally or log-normally distributed. In lieu of determining the nature of the distribution, USEPA suggests assuming that the data are log-normally distributed and use Land's formula to calculate an upper confidence limit for the mean (USEPA, 1992b). Land's formula uses the log-transformed data to calculate the mean and standard deviation of *n* sample data points and then calculates the upper confidence level of the sample data using the following equation (USEPA, 1992b):

$$\mathrm{UCL}_{1-\alpha} = \exp\left[\bar{x} + 0.5s_x^2 + \frac{s_x H_{1-\alpha}}{\sqrt{n-1}}\right]$$

Where $\bar{x}$ and s_x are the mean and standard deviation of the log-transformed data and $H_{1-\alpha}$ is the H-statistic for confidence level $1-\alpha$. USEPA recommends a minimum of four samples for estimating $\bar{x}$ and s_x however, tables for the H-statistic present values for sample sets with three samples (USEPA, 1992b).

The deterministic method has received much criticism for potentially over estimating the human health risk posed by a contaminated facility because of the use of overly conservative point estimates of risk model parameters. In addition, this method can result in "compounded conservatism". Bogen (1994) stated: " . . . *Compounded conservatism (or "creeping safety") describes the impact of using conservative upper-bound estimates of the values of multiple input variates to obtain a conservative estimate of risk modeled in an increasing function of those variates. In a simple multiplicative model of risk for example, if upper p-fractile (100pth percentile) values are used for each of several statistically independent input variates, the resulting risk estimate will be the upper p'-fractile of risk predicted according to that multiplicative model, where p' > p*."

Finley and Paustenbach, (1994a) stated that: "*A major shortcoming [of the deterministic approach] lies in the fact that repeated use of upper-bounded point estimates, as is recommended by the USEPA to calculate a reasonable maximal exposure (RME), typically leads to unrealistic estimates of health risk and unreasonable clean-up goals."* The issue of using overly conservative point estimates was addressed in two appendices to "Science and Judgment in Risk Assessment" (NRC, 1994). Appendix N-1 written by Adam M. Finkel and entitled "The Case for 'Plausible Conservatism' in Choosing and Altering Defaults" defended the USEPA approach to determining default values. He took the position that USEPA should evaluate its default parameters based on

two criteria, scientific plausibility and whether it is conservative and protective of human health (NRC, 1994). Appendix N-2 written by Roger O. McClellan and D. Warner North and entitled "Making Full Use of Scientific Information in Risk Assessment," advocates the use of "Plausible Conservatism" for choosing and altering default assumptions. Appendix N-2 suggests that more site-specific information should be incorporated into risk assessments to reduce possible default parameter conservatism (NRC, 1994).

There are several benefits of using point estimates of input variables for risk assessment. The procedure of risk assessment using point estimates is simple and accessible. Once contamination data have been collected applying the model of exposure assessment is very straightforward. Most inputs to the model can be found on the USEPA web site (http://www.epa.gov). In addition, regulators readily accept the risk assessment based on point estimates. This method has been used for more than 15 years, and there is a wealth of guidance for both conducting and reviewing deterministic risk assessments (USEPA, 1999a). Running the risk assessments in reverse can provide site concentration levels for a specified risk value. These levels can then be used to screen which site contaminants should be considered for further investigation (Finley and Paustenbach, 1994).

2.2.4 Probabilistic Risk Assessment

Probabilistic risk assessment is a generic term for risk assessments that use probability models for the input variables instead of point values. This approach makes possible quantitative analysis of variability and uncertainty associated with the input variables (USEPA, 1999a). The ability to quantitatively analyze variability and

uncertainty allows the risk manager to chose a regulatory level of concern that will be exceeded with some known probability in comparison with deterministic methods that provide an estimate of risk that represents a high end value with unknown probability (USEPA, 1999a).

Probabilistic risk assessment makes use of the same basic algorithms that were already described and used for deterministic risk assessment. However, the probability distribution of possible intake doses is determined before multiplying by the toxicity value. This results in a density distribution of risk values unlike the deterministic approach, which yields a single estimate that represents the risk associated with the RME (USEPA, 1999a). The USEPA does not recommend the use of probability distributions for the toxicity variable, because the data are not sufficient to generate distributions that represent its variability and uncertainty (USEPA, 1999a).

For the purposes of risk management decisions, the USEPA suggests using values that fall between the 90^{th} and 99.9^{th} percentile of the risk output distribution, termed as the RME range (USEPA, 1999a). The value that is chosen from this range should reflect the confidence in the input parameters to the probabilistic model, and the end use of the site being investigated, to ensure it is protective of human health and the environment (USEPA, 1999a).

The probabilistic approach to calculating risk has been thoroughly critiqued. Some of the disadvantages have been outlined by the USEPA (1999a). There may not be sufficient information about the variability and uncertainty to specify the appropriate distribution of the exposure variables. Using probabilistic methods require significantly more resources to perform the risk assessment, interpret the results, and present the final

report. USEPA also cautions against a false sense of accuracy in the risk estimates, since probabilistic methods only quantify variability and uncertainty, they are not reduced or eliminated.

Monte Carlo techniques are most frequently used to develop the input probability distributions needed for risk estimation. Burmaster and Anderson (1994) provide 14 principles of good practice for using Monte Carlo techniques in risk assessment. They reiterate USEPA's caution about use of inadequate data or inappropriate distributions for the input variables. Additionally, it is recommended that the input distributions be shown both numerically and graphically to convey the full sense of the distribution properties. Finally, they caution that enough iterations should be run to demonstrate that the tails of the simulated distribution stabilize.

Finley and Paustenbach (1994a) discussed the benefits of Monte Carlo-based probabilistic exposure assessment. They state that the deterministic method often predicts risks that are representative of the 99.9^{th} percentile for the range of values rather than the intended 50^{th} or 95^{th} percentiles. The USEPA points out that probabilistic methods can provide more insight into the characteristics of the entire risk distribution. Also, probabilistic methods allow for value-of-information analysis and sensitivity analysis for identifying the input variables that strongly influence the outcome of the probabilistic risk assessment (USEPA, 1999a).

The methods for conducting Monte Carlo-based risk assessments are outlined in "Risk Assessment Guidance for Superfund Volume III: Process of Conducting Probabilistic Risk Assessment (Part A)" (USEPA, 1999a). In Monte Carlo-based risk analysis, the upper-bound estimates of the variables used in the deterministic intake

model are replaced with probability distribution functions (pdfs) that describe the entire population of the particular variable (Dakins et al., 1996). A set of numbers for input to the intake model is obtained by randomly selecting a value from the pdf for each of the input variables. This set of values is then used to calculate one possible intake dose and an associated risk value. This process is iterated *n* times. The *n* possible values are combined to create pdfs of the intake dose and predicted risk. From the intake dose distribution, risk managers can select the level of conservatism (usually between the upper 90^{th} and 99.9^{th} percentiles of the distribution) that is appropriate to specify the reasonable maximum exposure (Dakins et al., 1996).

Random selections of values from the pdf can be performed using one of three possible sampling techniques (McKay et al., 1979). The first is random sampling where values are drawn at random from each of the input distributions. This objective is accomplished by using computer algorithms to generate series of pseudo-random numbers. These series are usually sufficiently long so that, for the purposes of risk assessment, they can be considered infinitely long (Press, 1986). The second method for selecting values from the input pdf is stratified sampling where the pdf is divided into N sections and a random sample is drawn from the N^{th} division of each pdf (McKay et al., 1979). The third method is Latin Hypercube Sampling, which is a subset of stratified sampling where each of the N sections of the pdf have the same marginal probability distribution (McKay et al., 1979). The latter two methods allow for faster convergence in the tails of the output distribution for the intake dose and risk (Burmaster and Anderson, 1994).

Monte Carlo-based risk assessment has been criticized because it is computationally more complicated and, therefore, more time consuming than deterministic risk assessment. Thousands of iterations are required so that the tails of the output pdf will stabilize, thus making quality control and calculation checking difficult. Even with this difficulty, USEPA has stated that "*Monte Carlo simulation is clearly superior to the qualitative procedures currently used to analyze uncertainty and variability*" (USEPA, 1997b) and has recently published guidance for conducting probabilistic risk assessments (USEPA, 1999a). The Virginia Department of Environment Quality has provided some guidance for performing probabilistic risk assessment in its Risk Exposure Analysis Management System (REAMS) guidance document (VADEQ, 1994).

Monte Carlo-based risk analysis may provide a more comprehensive estimation of environmental risk. By using the distribution rather than deterministic estimates of an input variable or parameter, variability and uncertainty can be accounted for quantitatively. By performing Monte Carlo-based risk analysis, risk managers can examine the full distribution of risk, and select a better reasonable maximum exposure value (USEPA, 1999a). Monte Carlo-based risk assessment obviates the arguments in the deterministic approach concerning the choice of a point estimate. It also eliminates the problem of compounded conservatism that arises from the multiplying several upper bound estimates or fiducial limits. Sensitivity analysis is more meaningful in conjunction with Monte Carlo-based analysis than with deterministic risk assessments (Finley and Paustenbach, 1994a).

Some of the drawbacks of the Monte Carlo-based approach can be overcome by using a combination of the probabilistic and deterministic methodology. In this approach, sensitivity analysis is first used to determine which input distributions can be treated as deterministic, based on their impact on the outcome of the model. Sensitivity analysis measures the relative impact that a variable has on the outcome of a model (Hamby, 1994), and can be used to determine which input variable distribution has the greatest impact on the calculated risk. Because inputs with greater variability and/or uncertainty have a greater impact on model outcomes in Monte Carlo assessments, sensitivity analysis can indicate which variables or parameters need further investigation and refinement. If an input parameter distribution has a relatively large impact on the model outcome, this indicates that its variability and/or uncertainty may be large. If the uncertainty is large, additional information may reduce this uncertainty and increase the accuracy of the risk estimate (Finley and Paustenbach, 1994a).

Generally, sensitivity analysis is performed by relating the change in a parameter to the corresponding change in the outcome of the model. Several measures of this relationship have been used in risk assessment. These are the Sensitivity Ratio, Pearson Correlation Coefficient, and Spearman Rank Order Correlation Coefficient (Palisade, 2000). The Sensitivity Ratio is the change in the output of a model divided by change in an input variable. Pearson Correlation Coefficient measures the strength and direction of the linear association between the values. Spearman Rank Order Correlation Coefficient is a nonparametric measure of the strength and direction of association of the ranks of two quantitative variables (USEPA, 1999a).

Once it has been determined which input distributions have negligible impact on the output distribution, these parameters can be treated as deterministic values (USEPA, 1999a). By treating these inputs as deterministic values, the number of pdfs is reduced. Thus the calculation checking problems and quality control problems are reduced, while the majority of the information concerning variability and uncertainty is retained in this combined probabilistic/deterministic approach.

Another approach that can make the probabilistic risk assessment methodology less computationally tedious is the "two-point" method (Guymon et al., 1981). However, it has never been used in environmental risk assessment. The two-point method is so named because it uses the first two moments of the random variables (X_1, X_2, . . .X_m) of a well-behaved multivariate distribution function $Y = f(X_1, X_2, \ldots X_m)$ to approximate the moments of Y (Rosenblueth, 1975). Only the mean and the standard deviation of each variable are needed instead of the entire probability density function. This reduces the required knowledge about the independent variables. Additionally, the two-point method reduces the number of iterations required to estimate an upper bound of the risk distribution. The two-point method has been applied to frost heave models (Guymon et al., 1981) and groundwater flow models (Guymon and Yen, 1990; Yen and Guymon, 1990)

Rosenblueth (1975) developed the two-point method. Consider X and Y are real random variables and Y is a well-behaved function of X, with $\overline{X}$ the expectation, σ_x the standard deviation, and ν_x the skewness coefficient of X. Rosenblueth (1975) defines the impulse probability density functions $f_+(X) = P_+\delta(X + x_+)$, $f_-(X) = P_-\delta(X - x_-)$, where P_+ and P_- are coefficients such that $P_+ + P_- = 1$, $\delta(x)$ is the

Dirac delta function and x_+ and x_- are arbitrary values of X greater than and less than $\overline{X}$ respectively. This implies that the n[th] moment $E(Y^n)$ of $Y = f_+(X) + f_-(X)$ is given by (Rosenblueth, 1975):

$$E(Y^n) = P_+ y_+^n + P_- y_-^n$$

where $y_+ = \delta(X + x_+)$, $y_- = \delta(X - x_-)$. This also implies P_+, P_-, x_+, x_- must satisfy the following simultaneous equations (Rosenblueth, 1975):

$$P_+ + P_- = 1$$

$$P_+ x_+ + P_- x_- = \overline{X}$$

$$P_+ \left(x_+ - \overline{X}\right)^2 + P_- (x_- - \overline{X})^2 = \sigma_x^2$$

$$P_+ \left(x_+ - \overline{X}\right)^3 + P_- (x_- - \overline{X})^3 = v_x^3 \sigma_x^3$$

whose solution is (Rosenblueth, 1975)

$$P_+ = \frac{1}{2}\left[1 \mp \sqrt{1 - \frac{1}{1 - (v_x/2)^2}}\right]$$

$$P_- = 1 - P_+$$

$$x_+ = \overline{X} + \sigma_x \sqrt{P_+ / P_-}$$

$$x_- = \overline{X} - \sigma_x \sqrt{P_- / P_+}$$

Here v_x is unknown and is assumed to be zero (Rosenblueth, 1975). This implies $P_+ = P_- = 1/2$ and $x_+ = \overline{X} + \sigma_x$, $x_- = \overline{X} - \sigma_x$. Assuming $v_x = 0$ restricts Y to a

symmetric, even function of X . The mean, standard deviation, and coefficient of variation for $Y = f_+(X) + f_-(X)$ is then given by (Rosenblueth, 1975):

$$\overline{Y} = \frac{Y_+ + Y_-}{2}$$

$$\sigma_Y = \left| \frac{Y_+ - Y_-}{2} \right|$$

$$V_Y = \left| \frac{Y_+ - Y_-}{Y_+ + Y_-} \right|$$

where $Y_+ = Y(x_+)$ and $Y_- = Y(x_-)$ and V_y is the coefficient of variation.

For more than one independent variables:.

$$Y = f(X_1, X_2, ...X_m) = f(X_1) \cdot f(X_2), ... \cdot f(X_m)$$

Assuming that none of these *m* variables are correlated, the n^{th} moment of Y can be calculated by:

$$E[Y^n] = \frac{1}{2^m}\left[(y'_{++...m})^n + (y'_{-+...m})^n + (y'_{--...m})^n\right]$$

where $E[Y^n]$ is the n^{th} moment of Y , m is the number of independent random variables, and the notation $y'_{++...m}$ indicates all permutations of the set:

$$y' = \{\overline{x}_1 \pm s_1, \overline{x}_2 \pm s_2, ...\overline{x}_m \pm s_m\}$$

where $\overline{x}_i$ is the mean, and s_i is the standard deviation of the i^{th} independent random variable (Guymon et al., 1981). From this relationship the mean of $Y = E(Y)$ is calculated when *n* is one and the variance as *VAR* [(Y)] = $E(Y^2)$ - $[E(Y)]^2$.

The two-point method requires 2^m iterations, which will be less than that required for the Monte Carlo-based simulation until the number of independent, random variables

exceeds 13. This method will also eliminate the problem of compound conservatism inherent in the deterministic method.

CHAPTER 3

MATERIALS AND METHODS

3.1 Brownfields Sites and Contaminant Data

In order to quantify the variability and uncertainty in risk estimates for individual contaminants using deterministic and probabilistic approaches a reliable set of contaminant data on brownfields sites were required. Risk estimates for multiple contaminants were not addressed in this study. This data set was developed using the Pennsylvania Department of Environmental Protection (PADEP) Land Recycling Program list of completed brownfields sites. This list is updated quarterly. The list at the end of the fall quarter of 2000 was used to develop the data set. The Fall 2000 list consisted of 633 completed brownfields. These sites were sorted into four groups based on whether the site was closed to background, statewide generic standards, site-specific, or industrial standards. Only the sites that were closed to statewide health, site-specific, and industrial standards were used in this study. From this reduced list, 30 sites were selected at random. The RAND function in EXCEL® (Microsoft Co., Redmond, WA) was used to generate 30 pseudo-random numbers. These numbers were then matched to specific sites on the numerically ordered reduced list. A request was make to PADEP to gain access to the files on these 30 sites. Ten of these sites did not have sufficient data or the PADEP could not find the files. The sites that were included in the study are presented in Table 1. One site, Exxon SGH Specialty Products, was listed as one facility but had two separate cleanup programs. Each of these was treated separately, giving a total of 21 brownfields sites used in this study.

The files on these 21 sites were archived in the specific PADEP Region listed in Table 1. This necessitated two separate trips to Pennsylvania to do the record reviews and collect data. Photocopies were made of all pertinent data in the files at the PADEP regional offices. The data were then manually transcribed into EXCEL® (Microsoft Co., Redmond, WA) spreadsheets. The data were carefully scrutinized for completeness and consistency. Since the PADEP has already approved these data for use in their risk assessments, and they had already been validated in accordance with USEPA guidelines as required under Pennsylvania's Land Recycling Program, a complete data validation was not needed.

Table 1. Twenty-one sites taken from Pennsylvania's Land Recycling Program list of completed brownfields.

Region	County	Town	Facility
Northeast	Lackawanna	Carbondale	Carbondale Railyards
Northwest	Lawrence	New Castle	Johnson Bronze
Northwest	Erie	Erie City	Erie City Iron Works
Southcentral	Blair	Hollidaysburg	GPU, Hollidaysburg Pole Yard
Southcentral	Cumberland	LeMoyne Borough	Conewago Contractors Inc.
Southcentral	Lebanon	West Lebanon	Aqua Chem Inc. Cleaver Brooks
Southeast	Bucks	Upper Makefield	Moyer Packing
Southeast	Bucks	Warrington	Home Depot
Southeast	Bucks	New Britain Township	Chalfont Plaza Assoc. LP
Southeast	Bucks	Perkasie Borough	LeNape MFG Co.
Southeast	Chester	Phoenixville Borough	West Co.
Southeast	Delaware	Chester	Peco Tighman St Gas PLT
Southeast	Delaware	Chester	SMK Speedy Intl. Inc.
Southeast	Delaware	Springfield Township	Sunoco 0004 8751
Southeast	Delaware	Upper Darby Township	Bond Shipping Center
Southeast	Philadelphia	Philadelphia	Van Waters and Rogers
Southeast	Philadelphia	Philadelphia	Flying Carport Inc.
Southeast	Philadelphia	Philadelphia	American Cable Co.
Southwest	Allegheny	Pittsburgh	Exxon SGH Specialty Products*
Southwest	Beaver	Freedom	Ashland Chemical Co.

* Two facilities listed as one because they belonged to the same company and were cleaned up as the same time.

Measured contaminant levels in the soil at these sites were used to estimate exposure concentrations in the air and groundwater. The concentrations were calculated using the procedures provided in the "Soil Screening Guidance: Technical Background Document" (USEPA, 1996). These procedures use several conservative assumptions to provide conservative estimates of air and groundwater concentrations. Most significantly, it is assumed that the soil is an unlimited source of the contaminant (USEPA, 1996). Also, the contaminant is assumed to be applied over a square area of half an acre that will provide conservative estimates for larger sites because of the unlimited source assumption. These assumptions simplify the underlying mass transfer models and reduce the need for large quantities of site-specific data that are often not available (USEPA, 1996).

The calculation of exposure concentrations for the inhalation route is divided into two possible mechanisms: inhalation of vapors for volatile contaminants and inhalation of contaminants adsorbed onto suspended soil particles for non-volatile contaminants. The concentration of contaminant in the air for volatile and non-volatile contaminants is calculated using the volatilization factor (VF) or the particulate emission factor (PEF), respectively. Both the VF and PEF are based on two models: one that estimates vapor or particulate emissions from the soil and one that estimates subsequent dispersion.

The emissions model for volatile contaminants is a simplification of the soil to atmosphere volatilization model developed by Jury et al. (1984) assuming that there is an unlimited source and no chemical and biological degradation. This model is based on transport theory outlined by Jury et al. (1983) and assumes 5 conditions. These are: (1) soil properties are constant with depth, (2) the contaminant partitioning on the soil

follows a linear, equilibrium adsorption isotherm, (3) partitioning of the contaminant between the soil solution and air follows Henry's law, (4) uniform initial concentration of contaminant in the soil, and (5) volatilization occurs through a stagnant air layer above which the contaminant concentration is zero (Jury et al., 1983). Additionally, it is assumed that the volatilization has reached steady state (USEPA, 1996).

The emissions model of the PEF is taken from Cowherd et al. (1985) who based his equation on the PM_{10} emission factor proposed by Gillette (1981). Calculation of the PEF requires that there is an unlimited reservoir of contaminated erodible particles. It is also assumed that the soils remain dry at all times, there is no variation in meteorology between the source and receptor, that the meteorology at the site can be represented as average values for the surrounding region, there is no deposition or reaction at the ground surface, and that emissions are uniformly distributed over the site. Additionally the level of contamination in the suspended particulate matter is assumed to be the same as the bulk soil (USEPA, 1996).

USEPA Office of Air Quality Planning and Standards developed the dispersion model for estimating ambient air concentrations from low or ground level, non-buoyant sources of emissions. The Industrial Source Complex Model platform was used to determine generic dispersion values for different regional meteorological conditions (USEPA, 1996). This dispersion model is integrated with either the vapor or particulate emission models resulting in single equations for estimating the VF and PEF. The procedures for estimating air concentrations from soil concentration are as follows:

Contaminant concentration in the air (USEPA, 1996)

$$CA = CS \times \frac{1}{VF} \quad \text{for volatile compounds}$$

$$CA = CS \times \frac{1}{PEF} \quad \text{for non-volatile compounds}$$

Where:

CA = Concentration of contaminant in the air at the soil surface (mg m^{-3})

CS = Concentration of contaminant in the soil (mg kg^{-1})

The volatilization factor is used to estimate the concentration of contaminant vapor in the air from volatile contaminants in the soil. The VF is calculated using chemical specific properties and default values of soil moisture, dry bulk density, and fraction of organic carbon, using the following relationship (USEPA, 1996):

$$VF = Q/C \times \frac{\left(3.14 \times D_A \times T\right)^{1/2}}{\left(2 \times \rho_b \times D_A\right)} \times 10^{-4}$$

VF = Soil to air volatilization factor
Q/C = Term linking atmospheric dispersion to contaminant vapor emissions from the soil surface and defined as the inverse of the mean concentration at the center of a ½ acre square source (g $m^{-2}s^{-1}$ (kg m^{-3})$^{-1}$)
T = Exposure intervals (s)
ρ_b = Dry soil bulk density (g cm^{-3})
D_A = Apparent diffusivity (cm^2 s^{-1})

Where:

$$D_A = \frac{\left[\left(\theta_a^{10/3} D_i H' + \theta_w^{10/3} D_w\right)/n^2\right]}{\rho_b K_d + \theta_w + \theta_a H'}$$

θ_a = Air-filled soil porosity
n = Total soil porosity
θ_w = Water filled soil porosity
D_i = Diffusivity in air ($cm^2\ s^{-1}$)
H' = Dimensionless Henry's law constant
D_w = Diffusivity in water ($cm^2\ s^{-1}$)
K_d = Soil water partition coefficient ($cm^3\ g^{-1}$)

The second possible mechanism for exposure to contaminants through the inhalation route is exposure to fugitive dust suspended by wind at the surface. The PEF is determined using the following relationship: (USEPA, 1996)

$$PEF = Q/C \times \frac{3600}{0.036 \times (1 - V) \times (U_m/U_t)^3 \times F(x)}$$

PEF = Particulate Emission Factor ($m^3\ kg^{-1}$)
Q/C = Term linking atmospheric dispersion to contaminant emissions from the soil surface and defined as the inverse of the mean concentration at the center of a ½ acre square source ($g\ m^{-2} s^{-1}\ (kg\ m^{-3})^{-1}$)
V = Fraction of vegetative cover (unitless)
U_m = Mean annual windspeed ($m\ s^{-1}$)
U_t = Equivalent threshold value of windspeed at 7 m ($m\ s^{-1}$)
$F(x)$ = Function dependent on U_m/U_t (unitless)
3600/0.36 = Empirical ratio ($m^2\ s\ g^{-1}$)

Parameter values required for calculating VF and PEF using the integrated soil to air emission and dispersion models are listed in Table 2.

Table 2 . Parameter values for the intergerated soil to air emission and dispersion models

Variable	Symbol	Value
Concentration of Contaminant in the air (mg m^{-3})	CA	Site Specific
Concentration of Contaminant in the soil (mg kg^{-1})	CS	Site Specific
Volatilization Factor	VF	Chemical Specific
Particulate Emission Factor (m^3 kg^{-1})	PEF	Site Specific
Atmospheric dispersion term (g m^{-2} (kg m^{-3})$^{-1}$)	Q/C	53.89
Exposure Intervals (s)	T	9.5 E8
dry soil bulk density (g cm^{-3})	ρ_b	1.5
Apparent diffusivity (cm^2 s^{-1})	D_A	Chemical Specific
Air-filled soil porosity for air	θ_a	0.28
Total soil porosity for air	n	0.43
Water filled soil porosity for air	θ_w	0.15
Diffusivity in air (cm^2 s^{-1})	D_i	Chemical Specific
Dimensionless Henry's law constant	H'	Chemical Specific
Diffusivity in water (cm^2 s^{-1})	D_w	Chemical Specific
Soil water partition coefficient (cm^3 g^{-1})	K_d	Chemical Specific
Fraction of vegetative cover (unitless)	V	0.5
Mean annual windspeed (m s^{-1})	U_m	4.69
Equivalent threshold value of windspeed at 7 m (m s^{-1})	U_t	11.32
Function dependent on U_m/U_t (unitless)	$F(x)$	0.194

The estimation of groundwater concentration depends on whether the contaminant is an organic or inorganic. The estimating equations are as follows (USEPA, 1996):

Inorganics

$$CW = CS \times \left(K_d + \frac{\theta_w + \theta_a H'}{\rho_b} \right)^{-1}$$

CS = Contaminant concentration in the soil (mg kg^{-1})
CW = Contaminant concentration in the soil leachate (mg L^{-1})
θ_a = Air-filled soil porosity
θ_w = Water filled soil porosity
H' = Dimensionless Henry's law constant
K_d = Soil water partition coefficient (cm^3 g^{-1})
ρ_b = Dry soil bulk density (g cm^{-3})

Organics

$$CW = CS \times \left(K_{OC} f_{OC} + \frac{\theta_w + \theta_a H'}{\rho_b} \right)^{-1}$$

CS = Contaminant concentration in the soil (mg kg^{-1})
θ_a = Air-filled soil porosity
θ_w = Water filled soil porosity
H' = Dimensionless Henry's law constant
K_{OC} = Soil water partition coefficient (cm^3 g^{-1})
f_{OC} = Fraction of organic carbon in soil
ρ_b = Dry soil bulk density (g cm^{-3})

These relationships hold for unlimited sources where the contaminants are uniformly distributed from the soil surface to the top of an unconsolidated, unconfined, homogenous, isotropic aquifer. It is also assumed that there is no attenuation by adsorption, chemical degradation, or biological degradation of the contaminants in the aquifer or the vadose zone, that there is instantaneous, linear soil/water partitioning in the vadose zone, that there are no non-aqueous phase liquids present, and that the receptor well is located at the edge of the source and is screened in the plume (USEPA, 1996). Parameters needed for estimating soil to ground water concentration are listed in Table 3 USEPA, 1996).

Table 3 . Parameter values for estimating groundwater concentraiton from soil concentration

Variable	Symbol	Value
Contaminant concentration in the groundwater (mg L^{-1})	CW	Site Specific
Soil water partition coefficient (cm^3 g^{-1})	K_{OC}	Chemical Specific
Fraction of organic carbon in soil	f_{OC}	0.002
Air-filled soil porosity	θ_a	0.30
Total soil porosity	n	0.43
Dimensionless Henry's law constant	H'	Chemical Specific
Water filled soil porosity	θ_w	0.13

The upper bound soil, air, and groundwater concentrations were calculated using Land's method as outlined "Supplemental Guidance to RAGS: Calculating the Concentration Term". Land's formula, as presented in Chapter 2, calculates the 95th percentile upper confidence limit (UCL) for the arithmetic mean of log transformed data, which is the default distribution assumed for most Superfund soil sampling data (USEPA, 1992b). When there is high standard deviation or low sample size Land's formula provides extremely high estimates for the UCL. In those cases the maximum detected or modeled concentration was used for the concentration term (USEPA, 1992b). Where the value of the sample concentration was reported as, non-detect or below the detection limit, the sample was treated as coming from a non-contaminated area and was removed from further consideration. Inclusion of samples with non-detect values is possible using the procedures outlined in "Chemical Concentration Data Near the Detection Limit" (USEPA, 1991). For several contaminant/site combinations the required information was not available in the files examined. Therefore, the procedures for treating non-detects were not feasible. By not including the non-detects as samples, the uncertainty in the 95th percentile UCL of the concentration is increased, thus increasing the standard deviation and increasing the concentration estimate (USEPA, 1991).

3.2 Deterministic Risk Estimates

The deterministic risk estimates for individual contaminants were calculated in accordance with the procedures outlined in "Risk Assessment Guidance for Superfund Volume I: Human Health Evaluation Manual (Part A)"(RAGS Vol. 1) Chapters 6, 7, and 8 (USEPA, 1989). As already discussed, the 95th percentile UCL soil concentration was calculated using Land's method as outlined in "Supplemental Guidance to RAGS:

Calculating the Concentration Term"(USEPA, 1991) and "Chemical Concentration Data Near the Detection Limit"(USEPA, 1992b). Ninety-fifth percentile UCL values of concentration used in calculating the deterministic risk estimations are contained in the Appendix. The air and groundwater concentrations were calculated using the algorithms provided in "Soil Screening Guidance: Technical Background Document" as outlined above. Other exposure parameters were taken from "The Exposure Assessment Handbook" and "USEPA Region III Risk-Based Concentration Table: Technical Background Information"(USEPA III, 1999) and are outlined with their source in Table 4.

The intake dose for the risk algorithm was calculated using the equations provided in Chapter 6 of RAGS Vol. 1 as previously noted in Chapter 2. The intake values were combined with the toxicity values obtained from the IRIS, HEAST, and ATSDR databases to calculate the deterministic risk estimates.

Table 4. Values of exposure factors for the deterministic and Monte Carlo/deterministic analyses

Exposure Variable	Symbol	Value	Source
Ingestion rate, adult (mg soil day^{-1})	IR	100	USEPA, 1999b
Fraction ingested from contaminated source (unitless)	FI	0.5	USEPA, 1999b
Exposure frequency (days $year^{-1}$)	EF	350	USEPA, 1999b
Exposure duration (years)	ED	30	USEPA, 1999b
Body weight (kg)	BW	70	USEPA, 1999b
Averaging time carcinogen (day)	AT	25550	USEPA, 1999b
Averaging time non-carcinogen (day)	AT	$ED \times 365$	USEPA, 1999b
Exposure time (hours day^{-1})	ET	24	USEPA, 1999b
Ingestion rate, tap water adult (L day^{-1})	IR	2	USEPA, 1999b
Skin surface area available for contact(cm^2)	SA	5800	USEPA, 1997
Soil to skin adherence factor (mg cm^{-2})	AF	0.55	USEPA, 1997
Inhalation rate, adult (m^3 day^{-1})	IR	20	USEPA, 1999b

3.3 Probabilistic Risk Estimates

The Monte Carlo and Monte Carlo/deterministic risk estimates for individual contaminants were calculated using @Risk 4.0, Advanced Risk Analysis for Spreadsheets (Palisade Corporation, Newfield, NY), a Monte Carlo simulator that is a plug-in to EXCEL® (Microsoft Co., Redmond, WA). Recommended probability density functions for the exposure variables were obtained from "Support Document for the Development of Generic Numerical Standards and Risk Assessment Procedures" issued by the Ohio Environmental Protection Agency (OEPA, 1996), and from Finley et al.(1994b). @Risk 4.0 calls specific function routines to generate the type of distribution that will be simulated. The arguments for the function routines are placed in parentheses after the function, and define the specific properties of the distribution. For the uniform distribution, minimum and maximum values define the range of possible values. The height and width of the triangular distribution are defined by minimum, most likely, and maximum values. The normal and lognormal distributions are defined by their mean and standard deviation. Both of these distributions can be truncated to include only a chosen range of values in the domain of the random variable. The general distribution is a probability density distribution based on a set of specified values of a random variable and their corresponding probability weights. The function parameters needed are a minimum and maximum value, a set of values, and a corresponding set of probability weights. A discrete uniform distribution has *n* possible outcomes of equal probability over an interval [a, b]. These values of a, b are the parameters of the function routine. The @Risk 4.0 function routines along with their arguments for the density distribution recommended for each exposure variable in the Monte Carlo and Monte Carlo/deterministic simulations are listed in Table 5.

Table 5. Recommended density distributions and @ Risk 4.0 function parameters (in parenthesis) for each exposure variable in the Monte Carlo and Monte Carlo/deterministic analysis

Exposure Variable	Distribution	Source
IR(mg soil day^{-1})	Uniform (min 10, max 200)	OEPA, 1996
FI(unitless)	Uniform (min 0.01, max 1)	OEPA, 1996
EF(days $year^{-1}$)	Triangular (261, 330, 365)	OEPA, 1996
ED(years)	General (min 0, max 70 (1,2,10,20,30,50), (0.18,0.27,0.13.0.2, 0.11, 0.11))	OEPA, 1996
BW(kg)	Normal (71, 15.9, Truncated at (32, 115))	OEPA, 1996
SA (cm^2)	Normal (284, 28, Truncated at (200, 351))	OEPA, 1996
FSA	Triangular (0.17, 0.42, 0.59)	OEPA, 1996
AF(mg cm^{-2})	Lognormal (0.52, 0.99, Truncated at (0.08, 3.44))	OEPA, 1996
IR(m^3air $hour^{-1}$)	Triangular (0.52, 0.8,1.02)	OEPA, 1996
ET(hours day^{-1})	Discrete uniform (16,24)	OEPA, 1996
IR(L water day^{-1})	General (min 0, max 5 (0.42,0.56,0.87,1.3,1.7,2.3,2.7,3.8,), (0.05,0.10,0.25,0.50,0.25,0.10,0.05,0.001))	Finley, 1994

Note: IR refers to soil ingestion, water ingestion, and air inhalation rates.

The probability density functions were fitted to the observed concentration data values using the @Risk 4.0 Best Fit function. The goodness of fit was determined using the Chi-Squared Statistic, the Kolmogorov-Smirnov statistic, and the Anderson-Darling statistic. Only the normal, lognormal, uniform, and triangular distributions were fitted to the soil concentration data. All of the best-fitted distributions are tabulated in the Appendix.

Ten thousand iterations of the Monte Carlo and Monte Carlo/Deterministic simulations were run using Latin Hypercube Sampling to ensure convergence for the tails of the output distribution. The results were then used to calculate the 90^{th}, 95^{th}, and 99.9^{th} percentiles of the intake dose distribution.

Sensitivity analysis was performed using the Monte Carlo simulations to determine which of the parameters could be treated as deterministic. The sensitivity analysis tool in @Risk 4.0 was used to perform the sensitivity analysis. Multivariate stepwise regression was used to determine the R-squared values. The parameters with the lowest R-squared values were then treated as deterministic values.

The risk estimates by the two-point method were calculated using computer algorithms coded in Matlab® (The Math Works Inc., Upper Saddle River, NJ). The required average and standard deviation for the soil concentrations were determined using EXCEL® (Microsoft Co., Redmond, WA) and are listed in the Appendix. The average and standard deviation for all other exposure variables were either collected from published data or determined from the distributions of the Monte Carlo simulations. They are listed in Table 6. Risk estimates for the two-point method were calculated for x_-, x_+ equal $\bar{x}$ minus or plus to one and two standard deviations.

Table 6. Mean and standard deviation of exposure variables used for the two-point-based risk estimation method

Exposure Variable	Average	Standard Deviation	Source
IR(mg soil day^{-1})	105	54.848	Monte Carlo
FI(unitless)	0.505	0.28579	Monte Carlo
EF(days year^{-1})	318.67	21.64	Monte Carlo
ED(years)	26.355	17.713	Monte Carlo
BW(kg)	71	15.9	Finley, 1994
SA (cm^2)	20181.58	4778.548	Monte Carlo
FSA	0.39333	0.086249	Monte Carlo
AF(mg cm^{-2})	0.52174	0.55302	Monte Carlo
IR(m^3 air hour^{-1})	0.78	0.10231	Monte Carlo
ET(hours day^{-1})	20	4	Monte Carlo
IR(L water day^{-1})	2.3609	0.80381	Finley, 1994

Note: IR refers to soil ingestion, water ingestion, or air inhalation rates.

3.4 Statistical Analysis

Statistical analyses were made to compare the estimated intake doses for the various estimation methods. Thirteen sets of intake doses were estimated and compared. These were the deterministic estimates and the 90^{th}, 95^{th}, and 99.9^{th} percentile of the estimated output distributions from the Monte Carlo, the Monte-Carlo/deterministic and the two-point method with x_-, x_+ equal to $\bar{x}$ minus or plus one and two times the standard deviation of each X. All of these estimated intake dose levels were laid out according to a randomized complete block design where each contaminant/site combination was a block and each estimating method was a treatment. An analysis of variance was made as appropriate for this design and the treatment means of the intake dose levels were compared for significant differences using Tukey's LSD. Analysis of the non-carcinogenic intake dose levels for the soil ingestion, soil dermal absorption, and groundwater ingestion exposure routes were performed for 69 contaminant/site combinations as listed in Table 7. Analysis of the non-carcinogenic inhalation intake dose levels was done for 45 contaminant/site combinations as listed in Table 7. Intake dose levels of carcinogens for 30 contaminant/site combinations were estimated for the soil ingestion, dermal absorption, and groundwater ingestion exposure routes and were analyzed as for the non-carcinogenic. These are listed in Table 8.

Table 7. Contaminant/site combinations and exposure routes for which intake dose for non-carcinogens were estimated

Contaminant	# of Sites	Soil Ingestion	Dermal Absorption	Inhalation	Groundwater Ingestion
Naphthalene	10	X	X	X	X
Ethylbenzene	7	X	X	X	X
Methylene chloride	7	X	X	X	X
Phenathrene	7	X	X	-	X
Toluene	6	X	X	X	X
Arsenic	6	X	X	-	X
Acetone	5	X	X	X	X
Barium	5	X	X	X	X
Chromium	5	X	X	X	X
Fluoranthene	5	X	X	-	X
Xylene	5	X	X	-	X
Totals	69	69	69	45	69

Table 8. Contaminant/site combinations and exposure routes for which intake dose for carcinogens were estimated

Contaminant	Site	Soil Ingestion	Dermal Absorption	Groundwater Ingestion
Arsenic	6	X	X	X
Benzo[a]anthracene	6	X	X	X
Benzo[a]pyrene	6	X	X	X
Chrysene	5	X	X	X
Methylene chloride	7	X	X	X
Totals	30	30	30	30

The calculated intake dose for carcinogens and non-carcinogens by each estimation method were separately sorted into groups by contaminant/site combinations and exposure route. This resulted in 13 separate sets of intake dose for each exposure route for both carcinogens and non-carcinogens. Linear regression analyses were made to determine the relationship between the intake dose levels of each set as the dependent variable and the standard deviation of the measured or modeled contaminant concentration for each contaminant/site combination as the independent variable. The

linear regression model was modified to determine if the response of the estimated intake dose to these standard deviations was the same for each of the 12 probabilistic methods as for the deterministic method. The modification uses dummy variables in the regression model as described by Ott (1993). The modified regression model is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + \varepsilon$$

where

y = the calculated exposure level

x_1 = standard deviation of the measured contaminant concentration for each contaminant/site combination

x_2 = the dummy variable with values 1 for the deterministic method and 0 for any of the probabilistic methods

β_0, β_1, β_2, and β_3, = regression coefficients

ε = the residual error assumed to be normally distributed with mean zero and standard deviation = 1.

For this model the expected value $E(y)$ of y for specified values of x_1, x_2 is given by:

$$E(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2$$

When the values of x_2 are substituted into the model the expected values for the deterministic and probabilistic method become (Ott, 1993):

Deterministic: $E(y) = \beta_0 + \beta_1 x_1 + \beta_2 + \beta_3 x_1 = (\beta_0 + \beta_2) + (\beta_1 + \beta_3) x_1$

Probabilistic: $E(y) = \beta_0 + \beta_1 x_1$

The fitted parameters β_2 and β_3 can be used to determine whether the regression line for the deterministic method and the corresponding line for any probabilistic method

are coincident, parallel, or equal. From the above equation for E(y), when β_2 is zero, the lines would have the same intercept and are said to be coincident. When β_3 is zero the lines have would the same slope and are said to be parallel. When β_2 and β_3 are both zero the lines are said to be equal. These criteria can be used to determine whether the intercept and slope of two regression lines are significantly different at a specified confidence level. All statistical analyses were performed using SAS (SAS Institute, Cary, NC).

CHAPTER 4

RESULTS AND DISCUSSION

Intake doses were generated for each contaminant/site combination by each method namely: the deterministic, the Monte Carlo probabilistic, the Monte Carlo probabilistic/deterministic, and the two point method using $x_+ = \overline{X} + s_x$ and $x_- = \overline{X} - s_x$ or $x_+ = \overline{X} + 2s_x$ and $x_- = \overline{X} - 2s_x$. The intake doses estimated by the deterministic method were fixed values. The probabilistic methods returned a distribution of intake doses. The intake dose for each of the probabilistic methods (the two point methods included) was taken as the 90^{th}, 95^{th} or the 99.9^{th} percentile of the estimated intake dose distribution. This resulted in 13 sets of intake doses for each of four exposure routes for carcinogens and three exposure routes for non-carcinogens. There were 69 contaminant/site combinations for the direct soil ingestion, dermal absorption, and groundwater ingestion exposure routes and 45 contaminant/site combinations for the inhalation exposure route for non-carcinogens. The direct soil ingestion, dermal absorption, and groundwater ingestion routes had 30 contaminant/site combinations for carcinogens. The inhalation exposure route for carcinogens did not have adequate contaminant/site data to perform reliable analyses and therefore was eliminated from further study. There were, therefore, a total of 897 intake dose values for the direct soil ingestion, dermal adsorption, and groundwater ingestion exposure routes for non-carcinogens, 585 values for the inhalation exposure route of non-carcinogens, and 390 values for each of the three exposure routes for carcinogens. Separate 1-way analyses of variance were made for carcinogens and non-carcinogens for each exposure route with the 13 methods as the treatments. Intake dose, instead of the actual risk estimates, were

used as the response variable of the 13 treatment methods to avoid the confounding effect due to differing levels of toxicity for the different chemicals. For ease of presentation, designations of the 13 treatments were abbreviated and these are listed in Table 9. These abbreviations will be used throughout the presentation of the results and discussion.

Table 9. Methods of risk estimation and their abbreviations. For each of the probabilistic methods the 90th, 95th or 99.9th percentile value of the estimated intake dose distribution was used for comparison

Method	Abbreviation
Deterministic	DET
Monte Carlo 90th Percentile	MC 90
Monte Carlo 95th Percentile	MC 95
Monte Carlo 99.9th Percentile	MC 99.9
Monte Carlo/deterministic 90th Percentile	MCD 90
Monte Carlo/deterministic 95th Percentile	MCD 95
Monte Carlo/deterministic 99.9th Percentile	MCD 99.9
Two-Point Method ($x_{\pm} = \overline{X} \pm s_x$) 90th Percentile	TP 90 (1)
Two-Point Method ($x_{\pm} = \overline{X} \pm s_x$) 95th Percentile	TP 95 (1)
Two-Point Method ($x_{\pm} = \overline{X} \pm s_x$) 99.9th Percentile	TP 99.9 (1)
Two-Point Method ($x_{\pm} = \overline{X} \pm 2s_x$) 90th Percentile	TP 90 (2)
Two-Point Method ($x_{\pm} = \overline{X} \pm 2s_x$) 95th Percentile	TP 95 (2)
Two-Point Method ($x_{\pm} = \overline{X} \pm 2s_x$) 99.9th Percentile	TP 99.9 (2)

Figure 1 shows the estimated intake dose averaged across all of the 69 contaminant/site combinations for each of the 13 estimation methods for the soil ingestion exposure route for non-carcinogens. The mean intake dose estimate was the key response variable used for comparing the various estimation methods in these analyses. In the remainder of this chapter it is abbreviated as the "MID estimate". The deterministic MID estimate was not significantly different from those for the MC 90 and

MC 95 or the MCD 90 and MCD 95 treatments. However, the 99.9th percentile MID estimates for the MC (Monte Carlo) and MCD (Monte Carlo/deterministic) methods were significantly different from the deterministic value. For all percentile levels of the MCD method, there were no significant differences in the MID estimates as compared to the corresponding MC generated values. The 90th and 95th percentile MID estimates calculated by the two-point methods were not significantly different from the corresponding values obtained by the MC and MCD methods. Nor were there significant differences between the MID estimate of the two-point methods at the 90th and 95th cutoff percentile and that of the deterministic method. The 99.9th percentile values obtained by the two-point methods were significantly different from the corresponding values for the MC and MCD methods. The MID estimate of the TP 99.9 (2) method was significantly different from the value estimated by the deterministic method. In summary, inclusion of deterministic values in the MCD estimation of the intake dose did not significantly influence the outcome compared to the MC method. Additionally, for all exposure routes, TP methods tended to give comparable MID estimates to the deterministic, MC, and MCD methods especially at the lower cutoff percentiles on the intake dose distribution.

Figure 2 shows the same MID estimates as in Figure 1 but for the dermal absorption exposure route. The MID estimates by this exposure route are an order of magnitude greater than those for the direct soil ingestion route. The pattern of the response to the various estimation methods is very similar to that in Figure 1. The MID estimate calculated using the deterministic method was not significantly different from

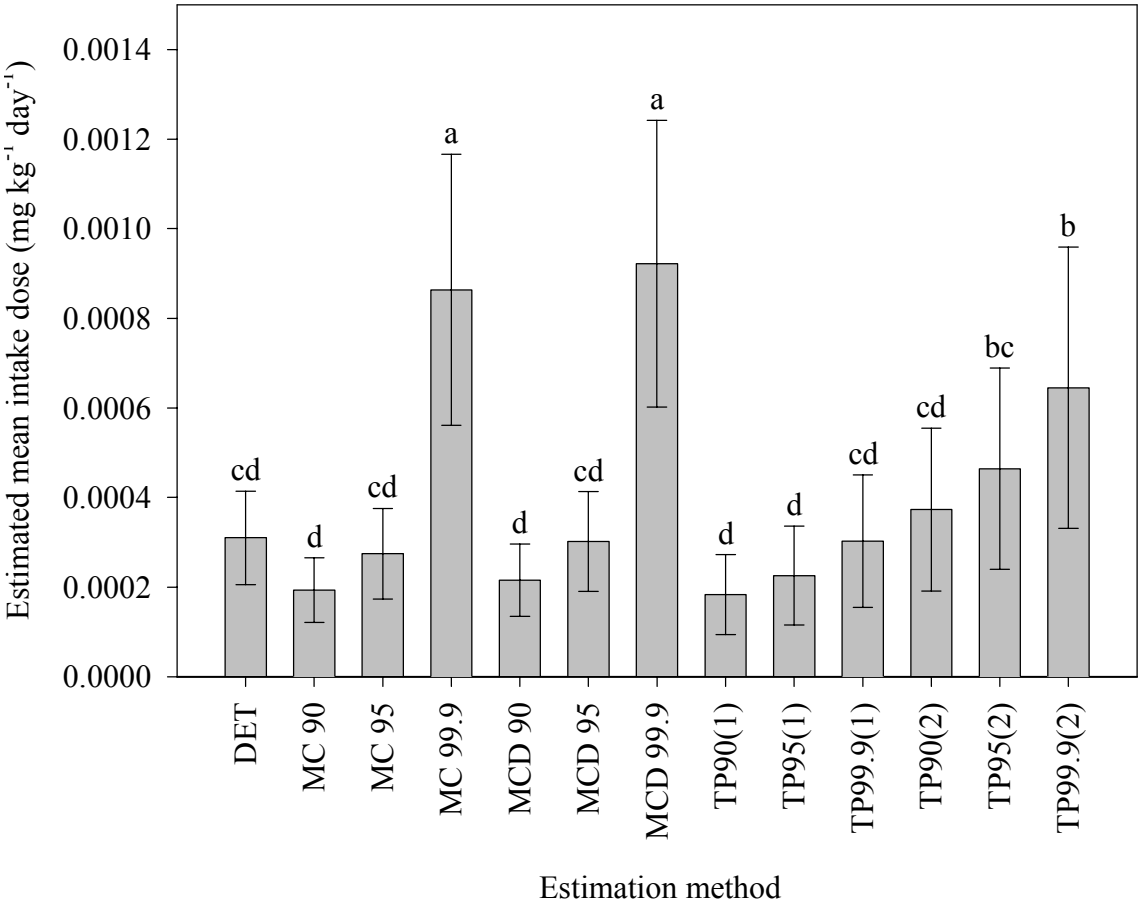

Figure 1. Mean estimated intake dose for the soil ingestion exposure route for non-carcinogens. Method abbreviations are defined in Table 9. Values not annotated with the same letter are significantly different at $p \leq 0.05$. Bars represent the standard error

those generated by the MC 90 and MC 95 or the MCD method at the 90th, 95th, and 99.9th percentiles. However, the MID estimate for the MC 99.9 method was significantly different from the value for the deterministic method. For all percentile levels of the MCD method there were no significant differences in the MID estimates compared to the corresponding values generated by the MC method. All of the two point methods generated MID estimates that were not significantly different from those generated by the deterministic method, the MC 90 and MC 95 methods, or all of the MCD methods. The MC 99.9 method generated significantly higher MID estimates than all the two-point methods.

Figure 3 shows the MID estimate over the 45 non-carcinogenic contaminant/site combinations for each of the 13 estimation methods for the inhalation exposure route. These MID estimates were respectively 3 and 2 orders of magnitude higher than those for non-carcinogens administered by the soil ingestion and dermal exposure routes. However, the pattern of the response of the MID estimates to the different methods is quite similar to those of the other two routes. The MID estimate calculated by the deterministic method was significantly different only when compared to the MC 99.9 and MCD 99.9 methods. MID estimates obtained using the two point methods were not significantly different from each other and from the MC and MCD methods at the 90th and 95th cutoff percentiles.

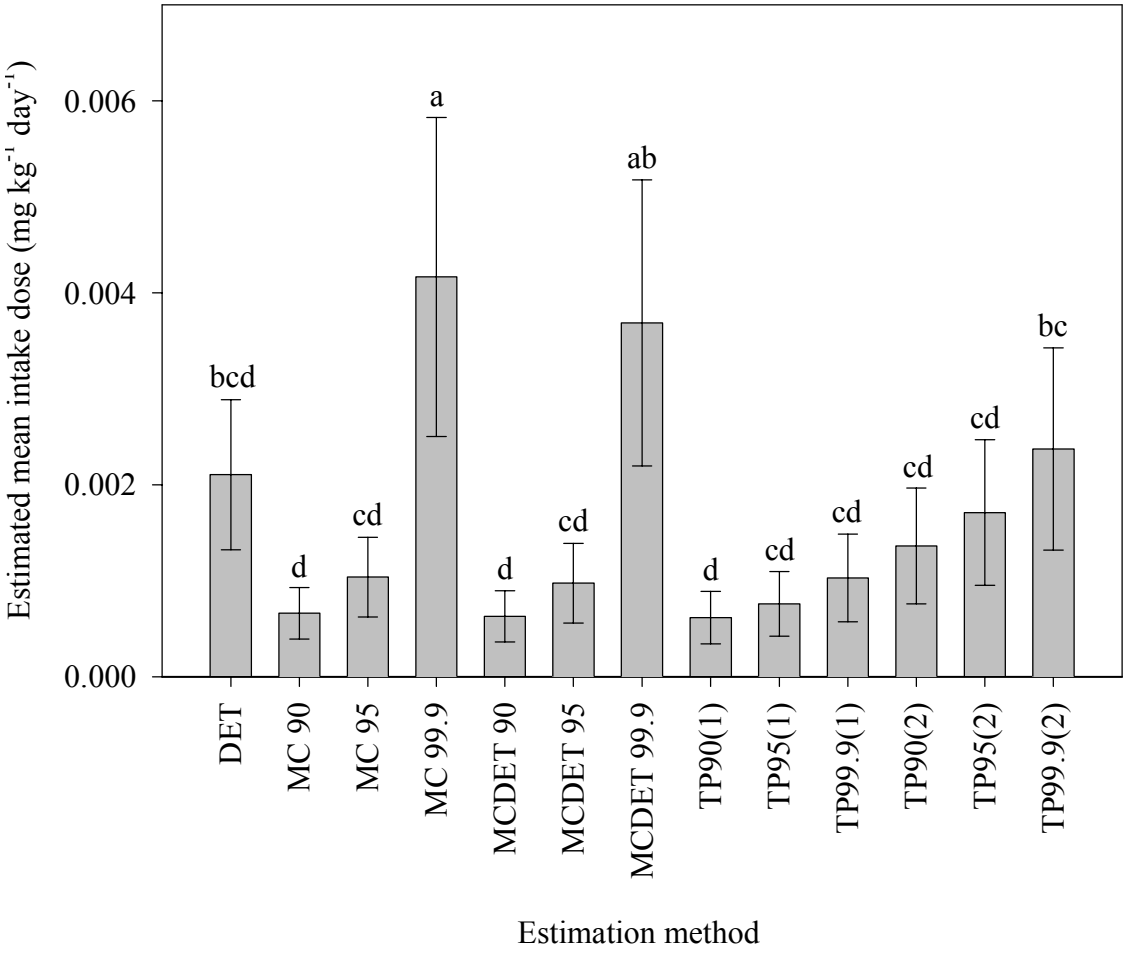

Figure 2. Mean estimated intake dose for the dermal exposure route for non-carcinogens. Method abbreviations are defined in Table 9. Values not annotated with the same letter are significantly different at $p \leq 0.05$. Bars represent the standard error

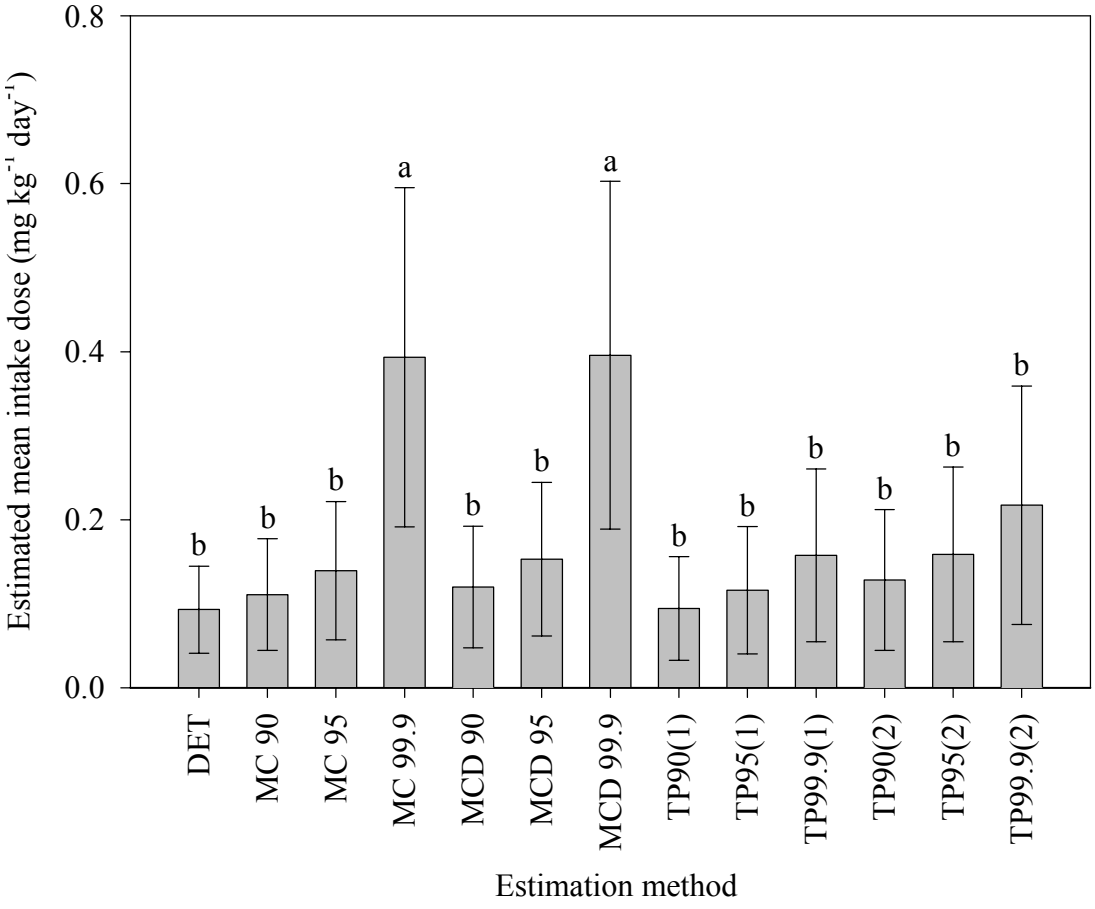

Figure 3. Mean estimated intake dose for the inhaletion exposure route for non-carcinogens. Method abbreviations are defined in Table 9. Values not annotated with the same letter are significantly different at $p \leq 0.05$. Bars represent the standard error

The MID estimates of the 69 contaminant/site combinations for each of the 13 estimation methods for the ingestion of groundwater exposure route for non-carcinogens are shown in Figure 4. These values were almost the same as for the inhalation exposure route except that the MID estimate of the deterministic method was 4 times greater. As a result, the deterministic MID estimate is not significantly different from those generated by the MC, MCD, and the TP 99.9 (2) methods. On the other hand, it was significantly higher than the MID estimates by all of the other methods. These latter MID estimates were not significantly different from each other.

Taken together the response of the MID estimates for non-carcinogens to the different estimation methods for the four exposure routes was very informative. The MID estimates for all methods were lowest for the direct soil ingestion exposure route. They increased by an order of magnitude for the dermal absorption route and 2 and 3 orders of magnitude for the inhalation and groundwater ingestion routes. This was not unexpected, since the contact rates for air and water were considerably higher than for soil ingestion and dermal absorption exposure routes as presented in Tables 4, 5, and 6. The much higher MID estimate by the deterministic method for the groundwater ingestion route versus the inhalation route was probably due to the different number of contaminant/site combinations (69 for the former versus 45 for the latter) used to compute the MID estimate. If this were the case, it may imply that the deterministic MID estimate is more sensitive to change in the level of variability or uncertainty associated with the contaminant/site data than the other estimation methods.

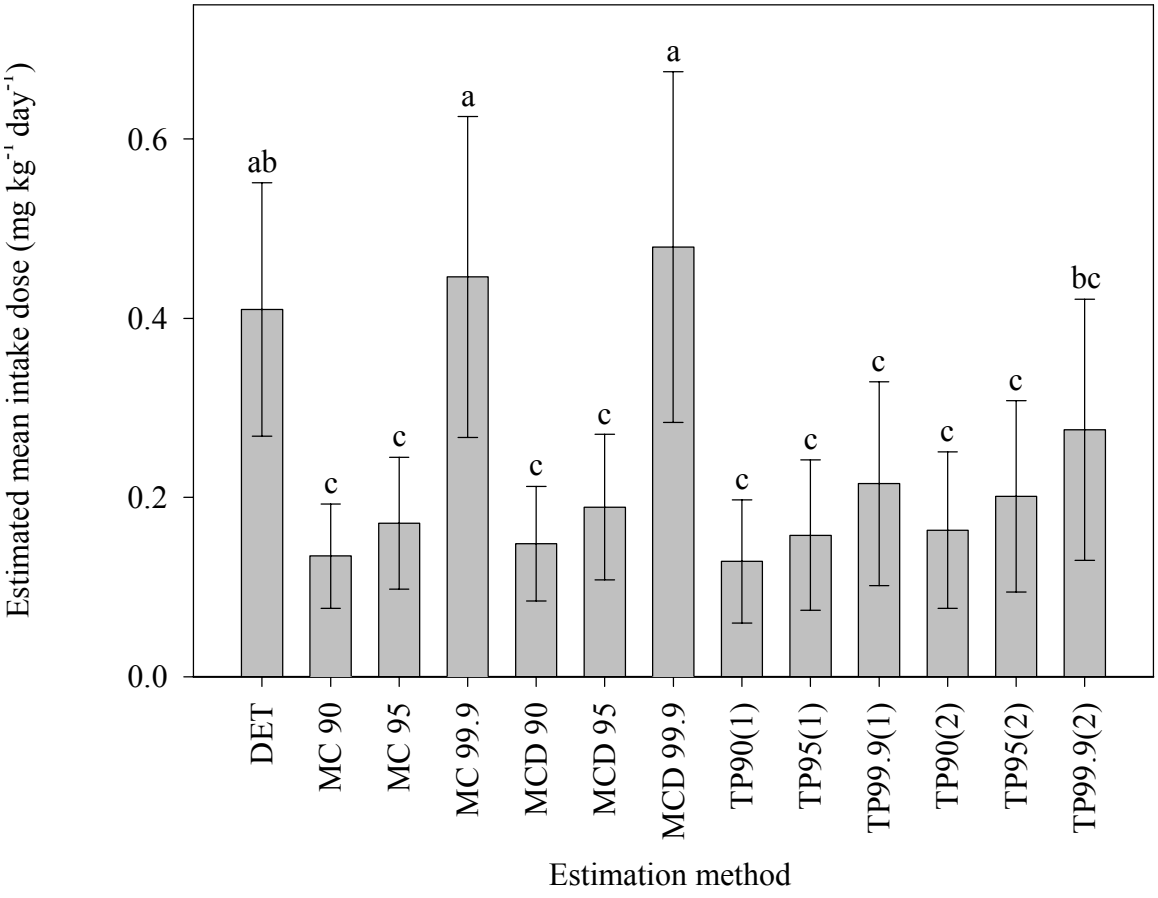

Figure 4. Mean estimated intake dose for the groundwater ingestion exposure route for non-carcinogens. Method abbreviations are defined in Table 9. Values not annotated with the same letter are significantly different at $p \leq 0.05$. Bars represent the standard error

Regardless of the exposure route the MID estimates showed a similar response pattern for the different estimation methods. As would be expected, the MID estimates for all the probabilistic methods increased non-linearly as the cutoff limits on the output probability distribution for the intake dose were raised from the 90^{th} to the 95^{th} and 99.9^{th} percentiles. However, this increasing trend was much more pronounced for the MC and MCD methods compared to the two-point methods for all exposure routes. This indicates that the tails of the output intake dose distribution generated by the two-point methods were relatively flatter compared to those for the MC and MCD methods. For all exposure routes the MC 99.9 and MCD 99.9 methods gave the highest MID estimates.

The risk estimates for non-carcinogens are calculated as the ratio of the intake dose divided by the RfD. These results therefore indicate that choosing the 99.9^{th} percentile as the cutoff limit for the MC and MCD methods would tend to yield much more conservative risk estimates that any of the other methods for all exposure routes. However, both of the two point methods, even at the 99.9^{th} percentile cutoff on the estimated intake dose distribution, would tend to yield risk estimates comparable to, or lower than the deterministic method, depending on the exposure route. The same is true when the MID estimates for the TP 99.9 (1) and TP 99.9 (2) methods were compared to those obtained by the MC and MCD methods with the 90^{th} and 95^{th} cutoff percentile on the output intake dose distribution. This is most likely due to the assumption of normality that is implicit in the development of the two-point method. The assumption of normality for all of the input variables would generate normally distributed intake dose estimates. As a result, the output distribution of the two-point method would tend to have less extended asymptotes than the MC or MCD methods. This effect of the assumption

of normality on the two-point output distributions appeared to be less pronounced in relation to the MC and MCD methods for all exposure routes at or below the 95th cutoff percentile. Consequently, this method may merit further examination in estimating risk for brownfields redevelopment, especially for cases, such as the redevelopment of industrial facilities for non-residential purposes, where using extreme upper-bound intake dose estimates are not warranted.

Figure 5 shows the MID estimates over the 30 carcinogenic contaminant/site combinations for each of the 13 estimation methods for the soil ingestion exposure route. Similar results are presented in Figures 6 and 7 for the dermal and groundwater ingestion exposure routes. As pointed out earlier, results for the inhalation exposure route for carcinogens were not generated since the available data set did not contain enough contaminant/site combinations nor toxicity values for carcinogens to adequately perform the analyses. As shown in Figures 5 to 7, the MID estimates are 1 or 2 orders of magnitude lower than those for non-carcinogens for a given exposure route. This is because the measured levels for carcinogens were relatively lower than those for non-carcinogens at all the 21 Brownfields sites (see Appendix). Additionally, for carcinogens, the averaging time is taken as 70 years, which is greater than the exposure duration whereas, for non-carcinogens it is the same as the exposure duration.

The MID estimates for carcinogens by the various methods follow the overall pattern shown in the results for non-carcinogens. The highest MID estimates were obtained with the MC 99.9 and MCD 99.9 methods, although, unlike the results for non-carcinogens, the latter were consistently lower than the former. As for non-carcinogens, the MID estimates for the carcinogens increased non-linearly as the cutoff limits on the

output intake dose distribution were raised to the higher percentiles. However, as shown in Figures 5 and 6, these increases were much steeper than those for non-carcinogens except for the groundwater ingestion route (Figure 7). This would indicate that the tails of the output intake dose distributions by the MC and MCD methods were much steeper than the corresponding ones for the non-carcinogens. This is probably caused by the lower variability in the measured concentrations for the carcinogens (see Appendix).

Viewed as a whole, the results of the analyses presented in Figures 1 though 7 were consistent with the findings of Smith (1994). The MID estimates obtained by the deterministic method for carcinogens and non-carcinogens for each exposure route were generally little different from those generated by the probabilistic methods except at the 99.9^{th} percentile cutoff on the output intake dose distribution. This finding does not appear to support the idea of compounded conservatism introduced by Bogen (1994) when he proposed that estimated intake dose and the associated risk is overestimated when using many upper-bound estimates in deterministic multiplicative models. This contradiction is only apparent, since several of the parameter values used in the deterministic computations in this study were not taken form the upper tails of their probability density distributions. As an example, the value for body weight (BW) was the mean of the body weight distribution. Use of such values would tend to reduce the compounded conservatism effect in the deterministic estimate.

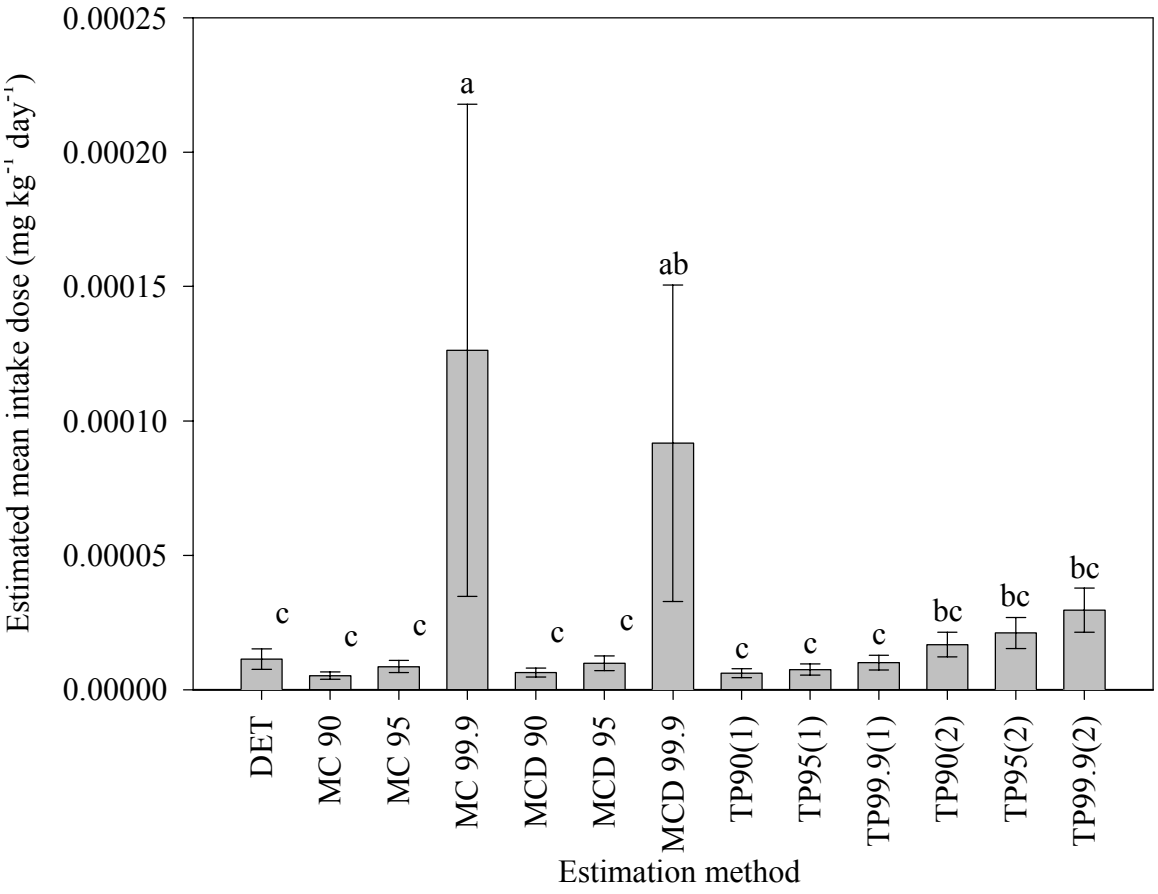

Figure 5. Mean estimated intake dose for the soil ingestion exposure route for carcinogens. Method abbreviations are defined in Table 9. Values not annotated with the same letter are significantly different at $p \leq 0.05$. Bars represent the standard error

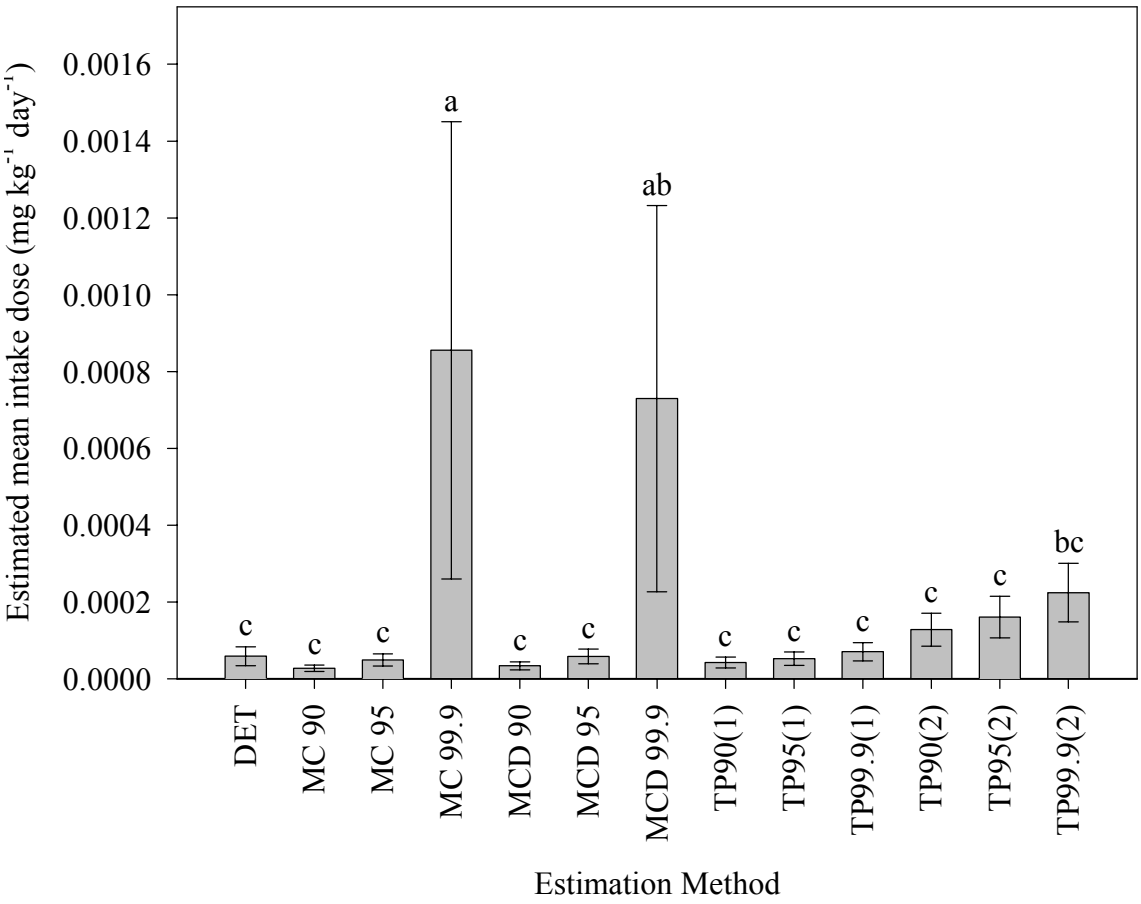

Figure 6. Mean estimated intake dose for the dermal exposure route carcinogens. Method abbreviations are defined in Table 9. Values not annotated with the same letter are significantly different at $p \leq 0.05$. Bars represent the standard error

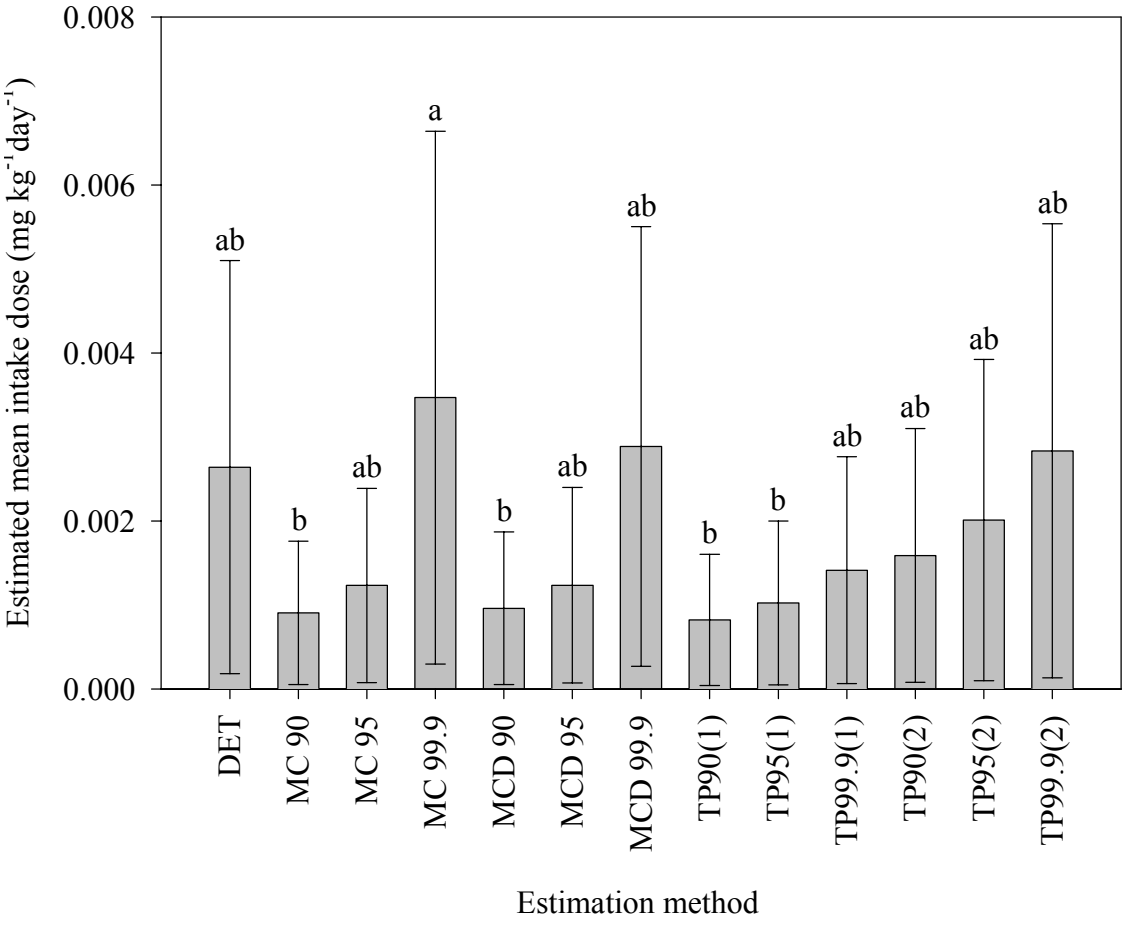

Figure 7. Mean estimated intake doses for the groundwater ingestion exposure route for carcinogens. Method abbreviations are defined in Table 9. Values not annotated with the same letter are significantly different at $p \leq 0.05$. Bars represent the standard error

The results in Figures 1 through 7 clearly show that for both carcinogens and non-carcinogens the MC and the MCD methods would tend to produce the same intake doses and risk estimates regardless of the exposure route. This is not surprising since using deterministic values in the Monte Carlo approach for the least sensitive input variables, as determined by sensitivity analyses, would not be expected to significantly impact the output intake dose distribution.

It should be noted that Figures 1 though 7 report mean intake dose averaged across all contaminant/site combinations for each method. This was done in order to evaluate trends in the relative performance of the estimation methods over varying contaminants, sites, and exposure routes. The foregoing results make a plausible case for using a tiered approach to determine the best method to assess the intake dose and associated risk at a given brownfields facility.

Further analyses were performed in an attempt to better evaluate how the variability and uncertainty in the measured soil contaminant concentrations at the 21 sites influenced the estimated intake doses by the deterministic vis-à-vis the probabilistic methods. The standard deviation of the concentration variable for each contaminant/site combination for each exposure route was calculated. These were then used as the independent variable in a modified linear regression (as described in Chapter 3) with the corresponding intake dose estimates as the dependent variable. Separate regressions were made for each of the estimation methods for each exposure route for carcinogens and non-carcinogens. A total of 91 such regressions were performed. The modified linear regression technique permitted a comparison of the intercept and slope of the regression line for the probabilistic methods for each exposure route with the corresponding values

for the deterministic method. However, this analytical approach cannot differentiate between variability and uncertainty in the standard deviation values of the contaminant concentrations at the brownfields sites. However, it was expected that general trends could be established and examined. The results of these analyses are presented in Tables 10 though 16.

Table 10 through 13 list the regression parameters for each of the 13 intake dose estimation methods generated using the values of the 69 contaminant/site combinations for each exposure route for non-carcinogens. Most of the regressions of the deterministic and Monte Carlo-based risk estimation method resulted in very weak correlation between the intake dose estimates and the standard deviation of the contaminant concentration. Several general trends are apparent in the results. All of the slopes of the regression lines are positive, indicating that a reduction in the contaminant concentration standard deviation may reduce the intake dose estimate. This indicates that additional sampling may be warranted in an attempt to reduce the standard deviation of the contaminant concentration. However, since the standard deviation encompasses all of the errors associated with the variability and uncertainty of the measured concentrations, additional sampling may not significantly reduce the standard deviation. Also, the cost of additional sampling must be weighed against the impact that the resulting risk estimate would have on reducing the redevelopment cost.

In general, the correlation between the intake dose estimates and the standard deviation of the contaminant concentration varied from one exposure route to another. The correlations were relatively good for the soil ingestion and groundwater ingestion routes compared to the dermal and inhalation routes. They were extremely poor for the

inhalation route. These findings were difficult to interpret. Linear least squares correlation is sensitive to a small number of high or low value pairs. It is possible that the different algorithms used for calculating the intake doses for the exposure routes can result in such pairs. This might explain the changing correlation from one exposure route to another. Except for the inhalation exposure route, the correlation for the probabilistic methods differed markedly for that for the deterministic method. This difference was less marked for the soil ingestion compared to the dermal absorption and groundwater ingestion exposure routes. Again these differences may be artifacts of the intake dose estimation algorithms for the deterministic versus the probabilistic methods. The correlation for the deterministic method was consistently poorer than those for the probabilistic methods, except for the inhalation exposure route where the r^2 values were comparable. This was not unexpected since the intake doses generated by the deterministic method were based on the 95^{th} percentile upper confidence level of the mean soil concentration.

For each exposure route, there were no significant differences between the regression parameters for the MC and MCD methods at the 95^{th} cutoff percentile as compared to the corresponding deterministic values. The slopes of the regression lines of the MC 90 method for soil ingestion, and the MCD 90 for dermal absorption were significantly lower than that of the deterministic method. A lower slope would indicate less influence of the variability and uncertainty in the concentrations. The slopes of the regression lines for the probabilistic method tended to increase as the cutoff percentile on the estimated intake dose distribution increased. In general these slopes were higher than that for the deterministic method especially at the 99.9^{th} cutoff percentile for all exposure

routes. This implied that the intake dose estimates would increase rapidly in response to variability and uncertainty in the sampled concentrations as the intake dose percentile cutoff was increased. This increase tended to be more pronounced for the MC and MCD methods that for the TP methods. Regardless for the exposure route, none of the regression intercepts for the Monte Carlo-based probabilistic methods were significantly different from those of the deterministic method. This indicated that the regression lines were coincident. Therefore, at low levels of variability and uncertainty the different estimation methods would not generate significantly different results. These results show that little advantage can be gained by using Monte Carlo-based methods as compared to the deterministic method if the intake dose estimate is selected in the low to moderate percentiles of the RME range.

The trends for the two-point method were similar to those of the Monte Carlo-based methods. For intake doses estimates in the low to moderate percentiles of the RME range, the two-point methods did not appear to provide an advantage over the deterministic method. However, intake dose estimates in the upper percentiles of the RME range for the two-point methods had a significantly greater regression slope that the deterministic method. For the soil ingestion exposure route, intake dose estimated by the TP 90 (1) and TP 95 (1) and by all the two-point methods for the groundwater exposure route had significantly different regression intercepts. This reinforces the conclusions from the ANOVA analysis for these two exposure routes that these methods provide significantly different intake dose estimates as compared with the deterministic method.

Table 10. Intercept, slope and coefficient of determination (r^2) for modified linear regression of intake doses by soil ingestion estimated by different methods, versus the standard deviation of the measured soil concentration of non-carcinogens

Estimation Method	Intercept x 10^{-5}	Slope x 10^{-6}	r^2
Deterministic	15.5	1.5	0.64
Monte Carlo 90	7.3	1.2*	0.80
Monte Carlo 95	10.6	1.7	0.80
Monte Carlo 99.9	36.3	5.0**	0.79
Monte Carlo/Deterministic 90	8.0	1.3	0.81
Monte Carlo/Deterministic 95	11.7	1.8	0.80
Monte Carlo/Deterministic 99.9	39.4	5.2**	0.79
Two Point 90 (1)	1.9*	1.6	0.98
Two Point 95 (1)	2.2*	2.0**	0.98
Two Point 99.9 (1)	3.1	2.7**	0.98
Two Point 90 (2)	3.9	3.3	0.98
Two Point 95 (2)	5.0	4.1	0.98
Two Point 99.9 (2)	6.7	5.7**	0.98

* and ** respectively indicate values are significantly different from the deterministic values at the 5% and 1% levels.

Table 11. Intercept, slope and coefficient of determination (r^2) for modified linear regression of intake doses by dermal absorption estimated by different methods, versus the standard deviation of the measured soil concentration levels of non-carcinogens

Estimation Method	Intercept x 10^{-4}	Slope x 10^{-6}	r^2
Deterministic	14.9	5.6	0.15
Monte Carlo 90	3.7	2.6	0.29
Monte Carlo 95	5.7	4.2	0.31
Monte Carlo 99.9	23.4	16.7**	0.30
Monte Carlo/Deterministic 90	4.7	1.4*	0.09
Monte Carlo/Deterministic 95	7.4	2.2	0.08
Monte Carlo/Deterministic 99.9	28.2	7.9	0.09
Two Point 90	1.1	4.6	0.86
Two Point 95	1.3	5.7	0.86
Two Point 99.9	1.8	7.8	0.86
Two Point (2) 90	2.4	10.3**	0.86
Two Point (2) 95	3.0	12.9**	0.86
Two Point (2) 99.9	4.1	17.9**	0.86

* and ** respectively indicate values are significantly different from the deterministic values at the 5% and 1% levels.

Table 12. Intercept, slope and coefficient of determination (r^2) for modified linear regression of intake doses by inhalation estimated by different methods, versus the standard deviation of the modeled air concentration of non-carcinogens

Estimation Method	Intercept x 10^{-3}	Slope	r^2
Deterministic	22.5	0.2	0.18
Monte Carlo 90	6.6	0.3	0.24
Monte Carlo 95	9.9	0.4	0.24
Monte Carlo 99.9	-25.3	1.4**	0.42
Monte Carlo/Deterministic 90	5.0	0.4	0.24
Monte Carlo/Deterministic 95	6.9	0.5	0.25
Monte Carlo/Deterministic 99.9	-29.1	1.4**	0.41
Two Point 90	4.0	0.3	0.21
Two Point 95	5.6	0.4	0.21
Two Point 99.9	8.6	0.5	0.20
Two Point (2) 90	5.6	0.4	0.21
Two Point (2) 95	7.7	0.5	0.21
Two Point (2) 99.9	11.7	0.7*	0.20

* and ** respectively indicate values are significantly different from the deterministic values at the 5% and 1% levels.

Table 13. Intercept, slope and coefficient of determination (r^2) for modified regression of intake doses by groundwater ingestion estimated by different methods, versus the standard deviation of the modeled groundwater concentration of non-carcinogens

Estimation Method	Intercept x 10^{-2}	Slope x 10^{-2}	r^2
Deterministic	27.9	6.5	0.25
Monte Carlo 90	3.6	4.9	0.84
Monte Carlo 95	4.7	6.1	0.83
Monte Carlo 99.9	41.5	14.7**	0.81
Monte Carlo/Deterministic 90	4.0	5.3	0.83
Monte Carlo/Deterministic 95	5.2	6.8	0.83
Monte Carlo/Deterministic 99.9	15.1	16.2**	0.82
Two Point 90	0.5*	6.1	0.95
Two Point 95	0.7*	7.5	0.95
Two Point 99.9	1.1*	10.1**	0.95
Two Point (2) 90	0.6*	7.8	0.95
Two Point (2) 95	0.9*	9.5*	0.95
Two Point (2) 99.9	1.4*	12.9**	0.95

* and ** respectively indicate values are significantly different from the deterministic values at the 5% and 1% levels.

Table 14 through 16 list the regression parameters for each of the 13 intake dose estimation methods generated using the values of 30 contaminant/site combinations for each exposure route for carcinogens. In the regression analyses for the groundwater ingestion route there was one extremely high standard deviation of the modeled groundwater concentration level. This was most probably an artifact of the soil to groundwater mass transfer algorithm. Therefore, this value was treated as an outlier and ignored for this set of analyses.

As for the non-carcinogens, the strength of the correlations between the intake dose estimates and the standard deviation of the input, concentrations depended with exposure route and estimation method. In addition, the correlations for the deterministic method were generally weaker than those of the probabilistic methods. They tended to be somewhat better for the soil ingestion and dermal exposure route compared to the groundwater ingestion route. They were uniformly high for the TP methods for all exposure routes and more variable for the Monte Carlo-based methods.

Interestingly, the correlation for the MC and MCD methods decreased as the cutoff percentiles on the output intake dose distribution were raised except for the groundwater ingestion exposure route. Like the non-carcinogens, the slopes of the regression lines increased as the cutoff percentile were increased for all the probabilistic methods and exposure routes.

The slopes of the regression lines for the probabilistic methods at the 90^{th} and 95^{th} cutoff percentiles were either lower than, or similar to, the slopes of the corresponding regression lines for the deterministic method. However, at the 99.9^{th} cutoff percentile the

slopes of the regression lines for all the probabilistic methods were significantly higher than the corresponding slopes for the deterministic method.

The regression parameters for the soil ingestion and dermal absorption exposure routes were similar. For the Monte Carlo-based probabilistic methods, the regression slope was significantly lower than that of the deterministic method when the intake dose estimates were taken at the 90^{th} percentile of the output probability distributions. However for these methods, the regression slope at the 95^{th} and 99.9^{th} percentile cutoff for the intake dose estimates was respectively not significantly different and significantly greater than that for the deterministic method. The regression slopes of the TP(2) methods were always significantly steeper than that of the deterministic method, whereas the TP(1) method at lower percentile cutoff for the RME had significantly flatter slopes than that of the deterministic method.

Table 14. Intercept, slope and coefficient of determination (r^2) for modified linear regression of intake doses by soil ingestion estimated by different methods, versus the standard deviation of the measured soil concentration levels of carcinogens

Estimation Method	Intercept x 10^{-6}	Slope x 10^{-6}	r^2
Deterministic	1.0	1.6	0.73
Monte Carlo 90	1.4	0.6**	0.80
Monte Carlo 95	1.8	1.0	0.89
Monte Carlo 99.9	-28.7	23.3**	0.28
Monte Carlo/Deterministic 90	1.5	0.7**	0.86
Monte Carlo/Deterministic 95	1.7	1.2	0.91
Monte Carlo/Deterministic 99.9	-13.6	15.8**	0.32
Two Point 90	1.0	0.8**	0.91
Two Point 95	1.2	0.9**	0.92
Two Point 99.9	1.5	1.3	0.93
Two Point (2) 90	2.8	2.1*	0.91
Two Point (2) 95	3.5	2.7**	0.92
Two Point (2) 99.9	4.5	3.8**	0.93

* and ** respectively indicate values are significantly different from the deterministic values at the 5% and 1% levels.

Table 15. Intercept, slope and coefficient of determination (r^2) for modified linear regression of intake doses by dermal absorption estimated by different methods, versus the standard deviation of the measured soil concentration of carcinogens.

Estimation Method	Intercept x 10^{-6}	Slope x 10^{-6}	r^2
Deterministic	-6.2	9.8	0.70
Monte Carlo 90	1.2	3.9**	0.90
Monte Carlo 95	0.3	7.3	0.94
Monte Carlo 99.9	-211.8	581.1**	0.32
Monte Carlo/Deterministic 90	1.4	3.8**	0.91
Monte Carlo/Deterministic 95	0.4	8.7	0.91
Monte Carlo/Deterministic 99.9	-176.1	136.1**	0.32
Two Point 90	-2.1	3.3*	0.96
Two Point 95	-2.9	4.0	0.96
Two Point 99.9	-4.4	11.6	0.96
Two Point (2) 90	-7.0	20.3**	0.96
Two Point (2) 95	-9.3	25.5**	0.97
Two Point (2) 99.9	-13.4	35.7**	0.97

* and ** respectively indicate values are significantly different from the deterministic values at the 5% and 1% levels.

Since an outlier value was ignored, Table 16 lists the regression parameters for each of the 13 intake dose estimation methods generated using the values of 29 contaminant/site combinations. For the Monte Carlo-based probabilistic methods, the regression slope was significantly steeper than that of the deterministic method when the intake dose estimates were taken at the 99.9^{th} percentile of the output probability distributions. All other regression parameters for all the other estimation methods were not different from those of the deterministic method. For4 this exposure route little or no advantage can be gained by reducing the standard deviation of the input concentration, when using Monte Carlo-based or two-point probabilistic methods as compared to deterministic methods if the intake dose estimate is selected in the low to moderate percentiles of the RME range.

Interpreting the results of the regression analyses taken as a whole were not straightforward. There were no consistent patterns in the linear relationship between the intake dose estimates and the standard deviations of the concentrations across exposure routes and estimation methods. Linear regression using log transformed intake dose estimates did not result in significant improvement in the correlations or in the separation of the regression parameters. The same was true using multivariate graphical techniques to uncover patterns in the response of the intake dose estimates to changes in the coefficient of variation of the measured or modeled concentrations. Techniques to analyze such dependence would merit further investigation.

Table 16. Intercept, slope and coefficient of determination (r^2) for regression of intake doses by groundwater ingestion estimated by different methods, versus the standard deviation of the modeled groundwater concentration of carcinogens.

Estimation Method	Intercept x 10^{-5}	Slope x 10^{-2}	r^2
Deterministic	9.1	9.6	0.14
Monte Carlo 90	1.9	3.5	0.43
Monte Carlo 95	2.4	5.2	0.48
Monte Carlo 99.9	0.5	30.3*	0.44
Monte Carlo/Deterministic 90	1.8	3.8	0.45
Monte Carlo/Deterministic 95	1.9	5.4	0.52
Monte Carlo/Deterministic 99.9	-2.2	30.8*	0.38
Two Point 90	0.6	3.6	0.85
Two Point 95	0.6	4.3	0.87
Two Point 99.9	0.8	5.7	0.89
Two Point (2) 90	1.1	7.0	0.84
Two Point (2) 95	1.4	8.7	0.85
Two Point (2) 99.9	1.8	11.9	0.86

* and ** respectively mean not significantly different, or significantly different at the 5% and 1% levels.

CHAPTER 5

SUMMARY AND CONCLUSIONS

Data on non-carcinogenic and carcinogenic contaminants at 21 brownfields sites in Pennsylvania were collected and compiled. Various methods were used to estimate intake doses and associated human risks for the principal environmental exposure routes of these contaminants. The results were used to determine whether the intake dose estimates for each exposure route depended on the estimation method and on the variability and uncertainty associated with the measured or modeled contaminant concentration at these sites. The commonly used deterministic method was evaluated along with three distinct probabilistic methods for estimating intake dose based (a) entirely on Monte Carlo simulated input parameter distribution functions, (b) on a combination of some of these functions and fixed parameter values, or (c) on a multivariate parameter distribution function. The deterministic method gave a fixed intake dose estimate for each contaminant/site combination. The probabilistic methods produced probability distributions for intake dose estimates. Consequently, the probabilistic methods permitted a choice of the intake dose for calculation of a risk value. Values corresponding to the 90^{th}, 95^{th} and 99.9^{th} percentiles of the intake dose distribution were used in this study. This resulted in a total of 13 sets of intake doses corresponding to 13 estimation methods.

Separate 1-way analyses of variance were made for carcinogens and non-carcinogens for each exposure route with the 13 methods as the treatments. Intake dose, instead of the actual risk estimates, was used as the response variable of the 13 treatment methods to avoid the confounding effect due to differing levels of toxicity for the

different chemicals. For non-carcinogens each set consisted of 69 contaminant/site combinations for the direct soil ingestion, dermal absorption, and groundwater ingestion exposure routes, and 45 contaminant/site combinations for the inhalation exposure route. The direct soil ingestion, dermal absorption, and groundwater ingestion routes had 30 contaminant/site combinations for carcinogens. There were, therefore, a total of 897 intake dose values for the direct soil ingestion, dermal adsorption, and groundwater ingestion exposure routes for non-carcinogens, 585 values for the inhalation exposure route of non-carcinogens, and 390 values for each of the three exposure routes for carcinogens.

The results of these analyses showed that use of probabilistic methods to estimate intake dose at the low to moderate cutoff levels (90^{th} and 95^{th} percentile) of the output distribution would not provide significantly different estimates as compared to deterministic method. For all intake dose cutoff percentile levels, and all exposure routes, the Monte Carlo/deterministic method generated intake dose estimates that were not significantly different from the corresponding values generated by the Monte Carlo method. There were no significant differences between the estimates generated by the two-point methods at the low and moderate percentile cutoff levels and the estimates generated by the deterministic or Monte Carlo-based methods. For the higher percentile cutoffs, the two-point method generally provided lower estimates of intake dose as compared to the Monte Carlo-based methods. This was most likely a result of the assumption of normality implicit in defining the multivariate distribution function for the two-point method.

Linear regression analyses were made to determine the correlation between each set of intake dose as the dependent variable and the standard deviation of the measured or modeled contaminant concentration as the independent variable. In these analyses the standard deviation was used as a measure of the variability and uncertainty in the contaminant concentration. The linear regression model was modified to determine if the response of the estimated intake dose to the standard deviation was the same for each of the 12 probabilistic methods as for the deterministic method. For all of the intake dose estimation methods, there was a weak positive linear relationship between the intake dose estimates and the standard deviation of the contaminant concentrations. The relationship was strongest for the intake dose estimates in the upper percentiles of the output distribution for the probabilistic methods. There was little difference in the strength and direction of the relationship for the deterministic method as compared to those for the probabilistic methods with intake dose chosen in the lower percentiles of the output distribution. Generally, the regression line intercepts for the probabilistic methods were not significantly different compared to the deterministic method. This indicated that at low standard deviation of contaminant concentration the methods would not generate significantly different intake dose estimates.

Taken together, these results indicated that there might be little advantage to choosing probabilistic methods over the deterministic method for risk assessment for brownfields redevelopment. However, for any particular brownfields facility, one of the probabilistic methods may provide more useful or useable estimates than the deterministic method. This lends support for the current tiered system of risk assessment. By increasing the level of sophistication at each tier, the likelihood of performing

unneeded analyses is reduced. The use of probabilistic risk assessment methods should be the last tier of such risk assessments.

References

Bogen, K.T. 1994. A note on compounded conservatism. Risk Analysis 14(4): 379-381.

Burmaster, D. E. and P. D. Anderson. 1994. Principles of good practice for the use of Monte Carlo techniques in human health and ecological risk assessments. Risk Analysis 14(4): 477-481

Byrd, D. M and C. R. Cothern. 2000. Introduction to Risk Analysis, A Systematic Approach to Science-Based Decision Making. Government Institutes. Rockville, Maryland.

Cowherd Jr., C. M., E. Gregory, P. J. Englehart, D. A. Gillette. 1985. Rapid assessment of exposure to particulate emissions from surface contamination sites. EPA/600/8/85/002, USEPA, Washington, D. C.

Dakins, M. E., J. E. Toll, M. J. Small, and K. P. Brand. 1996. Risk-based environmental remediation: Bayesian Monte Carlo analysis and the expected value of sample information. Risk Analysis 16(1): 67-79.

Davis, T. S. and K. D. Margolis. 1997. Brownfields: A comprehensive guide to redeveloping contaminated property. American Bar Association, Chicago, Illinois.

U. S. Environmental Protection Agency. 1989. Risk assessment guidance for Superfund volume I: human health evaluation manual (Part A). EPA/540/1-89/002, USEPA, Washington, D. C.

U. S. Environmental Protection Agency. 1991. Chemical concentration data near the detection limit. EPA/903/8-91/001, USEPA, Washington, D. C.

U. S. Environmental Protection Agency. 1992a. guidelines for exposure assessment. EPA/600/2-92/001, USEPA, Washington, D. C.

U. S. Environmental Protection Agency. 1992b. Supplemental guidance to RAGS: calculating the concentration term. USEPA, Washington, D. C.

U. S. Environmental Protection Agency. 1995. The brownfields action agenda. http://www.epa.gov/swerosps/bf/ascii/action.txt.

U. S. Environmental Protection Agency. 1996. Soil screening guidance: technical background document. EPA/540/R95/128, USEPA, Washington, D. C.

U. S. Environmental Protection Agency. 1997. Exposure factors handbook, volume 1-general factors. EPA/600/8-89/043, USEPA, Washington, D. C.

U. S. Environmental Protection Agency. 1997. Guiding principles for Monte Carlo analysis. EPA/630/R-97/001, USEPA, Washington, D. C.

U. S. Environmental Protection Agency. 1998. Brownfields assessment demonstration pilot, Shenandoah, VA. EPA/500/F/98/187, USEPA, Washington, D. C.

U. S. Environmental Protection Agency. 1999a. Risk assessment guidance for superfund volume 3: process for conduction probabilistic risk assessment (Part A)(Draft). USEPA, Washington, D. C.

U. S. Environmental Protection Agency. 1999b. brownfields assessment demonstration pilot, Newport News, VA. EPA/500/F/99/127, USEPA, Washington, D. C.

U. S. Environmental Protection Agency. 1999c. Office of solid waste and emergency response brownfields success stories: ecology in partnership with industry. http://www.epa.gov/swerosps/bf/html-doc/ss_cape.htm.

U. S. Environmental Protection Agency. 1999d. Office of solid waste and emergency response brownfields success stories: the right prescription: pharmaceutical company brings relief to Richmond. http://www.epa.gov/swerosps/bf/html-doc/ss_rich.htm.

U. S. Environmental Protection Agency. 2000a. Brownfields major milestones/accomplishments checklist. http://www.epa.gov/swerosps/bf/bfaccomp.htm

U. S. Environmental Protection Agency. 2000b. brownfields assessment demonstration pilot, Petersburg, VA.. EPA/500/F/00/114, USEPA, Washington, D. C.

U. S. Environmental Protection Agency Region III. 1999. USEPA region III risk-based concentration table: technical background information. USEPA, Washington, D. C.

Finley, B. and D. Paustenbachs. 1994a. The Benefits of Probabilistic Exposure Assessment: Three Case Studies Involving Contaminated Air, Water, and Soil. Risk Analysis 14(1): 53-73.

Finley, B., D. Proctor, P. Scott, N. Harrington, D. Paustenbach, and P. Price. 1994b. Recommended distributions for exposure factors frequently used in health risk assessment. Risk Analysis 14(4): 533-553.

Gillette , D. A. 1981. Production of dust that may be carried great distances. Geological Society of America Special Paper 186, 11-26.

Guymon, G. L., M. E. Harr, R. L. Berg, and T. V. Mromadka II . 1981. A probabilistic-deterministic analysis of one-dimensional ice segregation in a freezing soil column. Cold Regions Science and Technology 5(1981): 127-140.

Guymon, G. L. and C. Yen. 1990. An efficient deterministic-probabilistic approach to modeling regional groundwater flow 2. Application to Owens Valley, California. Water Resources Research 26(7): 1569-1581.

Hamby, D. M. 1994. A review of techniques for parameter sensitivity analysis of environmental models. Environmental Monitoring and Assessment 32: 135-154.

Jury, W. A., W. F. Spencer, and W. J. Farmer. 1983. Behavior assessment model for trace organics in soil: I. Model description. Journal of Environmental Quality 12(4), 558-564

Jury, W. A., W. J. Farmer, and W. F. Spencer. 1984. Behavior Assessment Model for Trace Organics in Soil: II. Chemical Classification and Parameter Sensitivity. Journal of Environmental Quality 13(4), 567-572

Land, C. E. 1971. Confidence intervals for linear functions of the normal mean and variance. The Annals of Mathematical Statistics 42(4), 1187-1205.

McKay, M. D., R. J. Beckman, and W. J. Conover. 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics 21(2): 239-244.

Morgan, M. G., and M. Henrion. 1990. Uncertainty. A guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, Cambridge, UK.

National Research Council (NRC). 1983. Risk assessment in the federal government: managing the process. Committee on the Institutional Means of Assessment of Risks to Public Health, Commission on Life Sciences, NRC. Washington, DC: National Academy Press.

National Research Council (NRC). 1994. Science and judgment in risk assessment. Washington, D.C, National Academy Press.

Ohio Environmental Protection Agency. 1996. Support document for the development of generic numerical standards and risk assessment procedures, the voluntary action program. OEPA, Columbus, Ohio.

Ott, R. L. 1993. An introduction to statistical methods and data analysis. Duxbury Press. Belmont, California.

Palisade. 2000. Guide to using @Risk, risk analysis and simulation add-in for Microsoft Excel. Palisade Corporation, Newfield, NY

Press, W. H. 1986. Numerical recipes in C: the art of scientific computing Cambridge; New York.

Rosenblueth, E. 1975. Point estimates for probability moments. Proceedings of the National Academy of Science, United States of America. 72(10): 3812-3814.

Rosenblueth, E. 1981. Two-point estimates in probabilities. Applied Mathematical Modeling 5: 329-335.

Smith, R. L. 1994. Use of Monte Carlo simulation for human exposure assessment at a Superfund site. Risk Analysis 14(4): 433-439.

Virginia Department of Environmental Quality. 1994. Guidance document and submission package for site remediation and cleanup using health based standards, Risk exposure and analysis modeling system [REAMS]. VADEQ, Richmond, Virginia.

Williams, P. L., R. C. James, and S. M. Roberts. 2000. Principles of toxicology: environmental and industrial applications. John Wiley & Sons, Inc., New York.

Yen, C., and G. L. Guymon. 1990. An efficient deterministic-probabilistic approach to modeling regional groundwater flow 1. Theory. Water Resources Research 26(7): 1559-1567

APPENDIX

Contaminant concentrations, their descriptive statistics and Land's UCL_{95}, calculated mass transfer parameters, and @Risk best fits for soil, air, and groundwater concentration distributions.

Appendix. Soil concentrations, descriptive statistics, Land's UCL_{95}, calculated mass transfer parameters, and @Risk best fits for soil, air, and groundwater

Contaminant (Type)	Site[1] (# of Samples)	Range	Mean	Std Dev	Median	Land's[2] UCL_{95}	VF[3]	PEF[4]	GW[5]	Distribution[6]		
		mg kg^{-1}								S[7]	A[8]	GW[9]
Naphthalene (non-carcinogen)	3 (4)	0.35-0.45	0.40	0.05	0.40	5.03E-01	N/A	1.42E+12	4.10E-03	TR	N/A	TR
Naphthalene (non-carcinogen)	5 (6)	0.425-1.99	1.08	0.56	1.05	2.48E+00	N/A	1.42E+12	4.10E-03	LN	N/A	N
Naphthalene (non-carcinogen)	6 (13)	0.13-0.23	0.17	0.06	0.14	2.03E-01	N/A	1.42E+12	4.10E-03	TR	N/A	TR
Naphthalene (non-carcinogen)	7 (7)	0.013-83.90	14.74	5.35	3.80	-	N/A	1.42E+12	4.10E-03	LN	N/A	N
Naphthalene (non-carcinogen)	8 (9)	0.23-2000	298.60	647.41	27.00	-	N/A	1.42E+12	4.10E-03	N	N/A	LN
Naphthalene (non-carcinogen)	10 (3)	3.9-18	9.33	7.59	6.10	3.39E+03	N/A	1.42E+12	4.10E-03	TR	N/A	TR
Naphthalene (non-carcinogen)	15 (3)	0.092-0.27	0.16	0.10	0.11	3.99E+00	N/A	1.42E+12	4.10E-03	TR	N/A	TR
Naphthalene (non-carcinogen)	16 (20)	0.68-200	16.23	43.78	3.80	3.80E+01	N/A	1.42E+12	4.10E-03	N	N/A	LN
Naphthalene (non-carcinogen)	17 (3)	0.03-0.69	0.36	0.33	0.37	-	N/A	1.42E+12	4.10E-03	TR	N/A	TR
Naphthalene (non-carcinogen)	19 (6)	0.01-0.013	0.01	0.001	0.01	1.20E-02	N/A	1.42E+12	4.10E-03	LN	N/A	N
Ethylbenzene (non-carcinogen)	7 (4)	1.8-9.70	4.05	3.79	2.35	4.76E+02	4.18E+03	N/A	2.08E-02	TR	TR	TR
Ethylbenzene (non-carcinogen)	8 (9)	0.023-3600	473.30	1191.31	0.56	-	4.18E+03	N/A	2.08E-02	N	N	N

Appendix. Continued

Contaminant (Type)	Site[1] (# of Samples)	Range	Mean	Std Dev	Median	Land's[2] UCL_{95}	VF[3]	PEF[4]	GW[5]	Distribution[6]		
		——— mg kg^{-1} ———								S[7]	A[8]	GW[9]
Ethylbenzene (non-carcinogen)	10 (4)	0.2-1.60	0.68	0.63	0.45	2.56E+02	4.18E+03	N/A	2.08E-02	TR	TR	TR
Ethylbenzene (non-carcinogen)	15(4)	0.006-79	19.76	39.50	0.01	-	4.18E+03	N/A	2.08E-02	TR	TR	TR
Ethylbenzene (non-carcinogen)	16 (16)	0.26-170	16.85	41.61	1.85	1.45E+02	4.18E+03	N/A	2.08E-02	LN	LN	LN
Ethylbenzene (non-carcinogen)	17 (6)	0.003-0.21	0.07	0.09	0.02	1.78E+02	4.18E+03	N/A	2.08E-02	N	LN	LN
Ethylbenzene (non-carcinogen)	18 (9)	0.02-0.70	0.20	0.22	0.14	1.38E+00	4.18E+03	N/A	2.08E-02	LN	LN	LN
Methylene Chloride (non-carcinogen/carcinogen)	3 (11)	0.0063-0.02	0.01	0.00	0.01	1.45E-02	1.93E+03	N/A	1.80E-01	N	N	N
Methylene Chloride (non-carcinogen/carcinogen)	8 (3)	0.0039-0.01	0.01	0.00	0.01	1.77E-01	1.93E+03	N/A	1.80E-01	TR	TR	TR
Methylene Chloride (non-carcinogen/carcinogen)	11 (7)	0.003-0.09	0.02	0.03	0.00	2.00E-01	1.93E+03	N/A	1.80E-01	LN	LN	N
Methylene Chloride (non-carcinogen/carcinogen)	13 (16)	0.0061-0.07	0.03	0.02	0.01	5.11E-02	1.93E+03	N/A	1.80E-01	U	U	LN
Methylene Chloride (non-carcinogen/carcinogen)	15 (8)	0.002-15	1.88	5.30	0.01	-	1.93E+03	N/A	1.80E-01	N	LN	N
Methylene Chloride (non-carcinogen/carcinogen)	20 (30)	0.005-0.06	0.01	0.01	0.01	1.11E-02	1.93E+03	N/A	1.80E-01	LN	LN	LN
Methylene Chloride (non-carcinogen/carcinogen)	21 (4)	0.0018-0.03	0.02	0.01	0.02	9.91E-02	1.93E+03	N/A	1.80E-01	TR	TR	TR
Phenanthrene (non-carcinogen)	3 (5)	0.4-1.30	0.68	0.39	0.47	1.64E+00	N/A	1.42E+12	1.72E-03	LN	LN	LN

Appendix. Continued

Contaminant (Type)	Site[1] (# of Samples)	Range	Mean	Std Dev	Median	Land's[2] UCL_{95}	VF[3]	PEF[4]	GW[5]	Distribution[6]		
		mg kg^{-1}								S[7]	A[8]	GW[9]
Phenanthrene (non-carcinogen)	5 (10)	0.364-66	25.58	22.84	20.25	7.15E+02	N/A	1.42E+12	1.72E-03	LN	LN	LN
Phenanthrene (non-carcinogen)	6 (3)	0.12-0.93	0.43	0.29	0.35	3.46E+01	N/A	1.42E+12	1.72E-03	LN	LN	LN
Phenanthrene (non-carcinogen)	8 (14)	0.15-930	164.26	281.11	9.85	-	N/A	1.42E+12	1.72E-03	N	N	N
Phenanthrene (non-carcinogen)	10 (4)	22007.00	22.68	25.55	13.35	-	N/A	1.42E+12	1.72E-03	TR	TR	TR
Phenanthrene (non-carcinogen)	15 (4)	0.18-0.31	0.24	0.06	0.23	4.50E-01	N/A	1.42E+12	1.72E-03	TR	TR	TR
Phenanthrene (non-carcinogen)	17 (11)	0.01-1.48	0.51	0.52	0.16	1.04E+01	N/A	1.42E+12	1.72E-03	LN	LN	LN
Toluene (non-carcinogen)	6 (5)	0.14-0.46	0.25	0.14	0.18	5.99E-01	3.08E+03	N/A	3.80E-02	N	LN	LN
Toluene (non-carcinogen)	8 (14)	0.004-2100	150.20	561.19	0.04	-	3.08E+03	N/A	3.80E-02	LN	LN	N
Toluene (non-carcinogen)	15 (3)	0.015-1.40	0.48	0.80	0.02	-	3.08E+03	N/A	3.80E-02	TR	TR	TR
Toluene (non-carcinogen)	16 (8)	0.21-1.20	0.45	0.34	0.31	8.69E-01	3.08E+03	N/A	3.80E-02	LN	LN	N
Toluene (non-carcinogen)	17 (8)	0.003-0.16	0.03	0.05	0.01	3.17E-01	3.08E+03	N/A	3.80E-02	LN	LN	LN
Toluene (non-carcinogen)	18 (15)	0.04-163	10.17	34.54	0.36	2.93E+02	3.08E+03	N/A	3.80E-02	LN	N	LN

Appendix. Continued

Contaminant (Type)	Site[1] (# of Samples)	Range	Mean	Std Dev	Median	Land's[2] UCL_{95}	VF[3]	PEF[4]	GW[5]	Distribution[6]		
				mg kg^{-1}						S[7]	A[8]	GW[9]
Toluene (non-carcinogen)	21 (6)	0.008-0.08	0.03	0.03	0.02	1.51E-01	3.08E+03	N/A	3.80E-02	LN	N	N
Arsenic (non-carcinogen/carcinogen)	3 (9)	2.5-32	8.82	9.45	5.30	2.28E+01	N/A	1.42E+12	2.50E-04	LN	N/A	LN
Arsenic (non-carcinogen/carcinogen)	4 (15)	3.3-28.8	15.99	8.03	16.10	2.55E+01	N/A	1.42E+12	2.50E-04	N	N/A	N
Arsenic (non-carcinogen/carcinogen)	5 (32)	14.2-59.20	30.16	13.04	27.75	3.52E+01	N/A	1.42E+12	2.50E-04	LN	N/A	LN
Arsenic (non-carcinogen/carcinogen)	6 (24)	2.6-20	7.76	4.35	6.75	9.79E+00	N/A	1.42E+12	2.50E-04	LN	N/A	TR
Arsenic (non-carcinogen/carcinogen)	9 (3)	5.5-9.59	7.52	2.05	7.46	1.71E+01	N/A	1.42E+12	2.50E-04	TR	N/A	TR
Arsenic (non-carcinogen/carcinogen)	20 (25)	0.6-10.50	3.16	2.99	1.90	5.44E+00	N/A	1.42E+12	2.50E-04	N	N/A	N
Benzo[a]anthracene (carcinogen)	3 (4)	0.34-0.65	0.51	0.14	0.53	1.04E+00	N/A	1.42E+12	2.09E-06	TR	N/A	TR
Benzo[a]anthracene (carcinogen)	5 (9)	1.89-46.00	18.96	15.43	15.20	1.46E+02	N/A	1.42E+12	2.09E-06	LN	N/A	LN
Benzo[a]anthracene (carcinogen)	6 (16)	0.078-1.10	0.31	0.31	0.20	5.36E-01	N/A	1.42E+12	2.09E-06	LN	N/A	LN
Benzo[a]anthracene (carcinogen)	8 (16)	0.0015-130	17.26	32.27	2.75	-	N/A	1.42E+12	2.09E-06	N	N/A	LN
Benzo[a]anthracene (carcinogen)	15 (4)	0.15-0.27	0.21	0.05	0.21	3.86E-01	N/A	1.42E+12	2.09E-06	TR	N/A	TR
Benzo[a]anthracene (carcinogen)	17 (14)	0.006-0.69	0.21	0.23	0.11	1.96E+00	N/A	1.42E+12	2.09E-06	LN	N/A	N

Appendix. Continued

Contaminant (Type)	Site[1] (# of Samples)	Range	Mean	Std Dev	Median	Land's[2] UCL_{95}	VF[3]	PEF[4]	GW[5]	Distribution[6]		
		——— mg kg^{-1} ———								S[7]	A[8]	GW[9]
Benzo[a]pyrene (carcinogen)	3 (4)	0.39-1	0.64	0.27	0.58	3.03E+00	N/A	1.42E+12	8.17E-06	TR	N/A	TR
Benzo[a]pyrene (carcinogen)	5 (9)	2.19-50.50	16.25	16.96	15.00	-	N/A	1.42E+12	8.17E-06	LN	N/A	LN
Benzo[a]pyrene (carcinogen)	6 (10)	0.075-0.71	0.31	0.22	0.25	6.13E-01	N/A	1.42E+12	8.17E-06	LN	N/A	LN
Benzo[a]pyrene (carcinogen)	8 (15)	0.0011-150	25.17	44.24	8.50	-	N/A	1.42E+12	8.17E-06	LN	N/A	LN
Benzo[a]pyrene (carcinogen)	9 (11)	0.94-30.20	6.21	8.63	3.08	1.56E+01	N/A	1.42E+12	8.17E-06	LN	N/A	LN
Benzo[a]pyrene (carcinogen)	17 (17)	0.004-1.03	0.24	0.31	0.10	1.55E+00	N/A	1.42E+12	8.17E-06	N	N/A	LN
Acetone (non-carcinogen)	1 (3)	0.039-9.30	3.15	5.33	0.10	-	9.72E+03	N/A	2.46E-01	TR	TR	TR
Acetone (non-carcinogen)	8 (14)	0.0086-350	25.20	93.48	0.12	8.13E+02	9.72E+03	N/A	2.46E-01	N	N	LN
Acetone (non-carcinogen)	11 (10)	0.016-5.90	0.65	1.85	0.03	8.95E+00	9.72E+03	N/A	2.46E-01	N	LN	LN
Acetone (non-carcinogen)	15 (7)	0.026-0.11	0.06	0.03	0.05	1.11E-01	9.72E+03	N/A	2.46E-01	LN	LN	LN
Acetone (non-carcinogen)	17 (12)	0.105-13.84	2.30	3.80	0.60	2.57E+01	9.72E+03	N/A	2.46E-01	N	LN	N
Barium (non-carcinogen)	3 (10)	24-88	43.40	23.21	35.00	6.26E+01	N/A	1.42E+12	8.31E-04	N	LN	N

Appendix. Continued

Contaminant (Type)	Site[1] (# of Samples)	Range	Mean	Std Dev	Median	Land's[2] UCL_{95}	VF[3]	PEF[4]	GW[5]	Distribution[6]		
		mg kg^{-1}								S[7]	A[8]	GW[9]
Barium (non-carcinogen)	4 (15)	30-256	125.27	59.99	125.00	2.14E+02	N/A	1.42E+12	8.31E-04	LN	LN	LN
Barium (non-carcinogen)	5 (32)	24.5-1640	335.09	347.44	200.00	5.27E+02	N/A	1.42E+12	8.31E-04	LN	TR	LN
Barium (non-carcinogen)	6 (24)	15-80	40.46	14.32	40.00	4.78E+01	N/A	1.42E+12	8.31E-04	TR	TR	TR
Barium (non-carcinogen)	20 (30)	48-890	193.40	194.96	135.00	2.40E+02	N/A	1.42E+12	8.31E-04	LN	LN	LN
Chromium (non-carcinogen)	3 (8)	7.8-15	12.64	2.87	14.00	1.56E+01	N/A	1.42E+12	1.34E-03	N	N/A	N
Chromium (non-carcinogen)	4 (12)	5-326	72.00	100.17	23.00	3.79E+02	N/A	1.42E+12	1.34E-03	LN	N/A	LN
Chromium (non-carcinogen)	5 (32)	6.61-75	16.33	12.54	11.85	1.90E+01	N/A	1.42E+12	1.34E-03	LN	N/A	N
Chromium (non-carcinogen)	6 (24)	2.4-14	7.30	2.53	6.90	8.51E+00	N/A	1.42E+12	1.34E-03	N	N/A	N
Chromium (non-carcinogen)	20 (30)	12693.00	20.10	4.73	20.00	2.19E+01	N/A	1.42E+12	1.34E-03	N	N/A	N
Chrysene (carcinogen)	3 (4)	0.5-0.87	0.70	0.17	0.72	1.30E+00	N/A	1.42E+12	2.09E-05	TR	N/A	TR
Chrysene (carcinogen)	6 (8)	0.14-1.70	0.52	0.49	0.32	1.52E+00	N/A	1.42E+12	2.09E-05	N	N/A	LN
Chrysene (carcinogen)	8 (14)	0.15-130	17.12	33.99	3.45	3.33E+02	N/A	1.42E+12	2.09E-05	LN	N/A	N
Chrysene (carcinogen)	15 (4)	0.11-0.48	0.31	0.18	0.32	1.26E+01	N/A	1.42E+12	2.09E-05	TR	N/A	TR

Appendix. Continued

Contaminant (Type)	Site[1] (# of Samples)	Range	Mean	Std Dev	Median	Land's[2] UCL_{95}	VF[3]	PEF[4]	GW[5]	Distribution[6]		
		mg kg^{-1}								S[7]	A[8]	GW[9]
Chrysene (carcinogen)	17 (13)	0.006-1.09	0.30	0.36	0.13	3.52E+00	N/A	1.42E+12	2.09E-05	LN	N/A	LN
Fluoranthene (non-carcinogen)	3 (4)	0.45-1.50	0.87	0.45	0.76	5.81E+00	N/A	1.42E+12	7.79E-05	TR	N/A	TR
Fluoranthene (non-carcinogen)	6 (3)	0.13-2.20	0.61	0.60	0.42	8.49E+02	N/A	1.42E+12	7.79E-05	LN	N/A	LN
Fluoranthene (non-carcinogen)	8 (14)	0.11-250	41.28	66.27	20.70	5.06E+03	N/A	1.42E+12	7.79E-05	N	N/A	LN
Fluoranthene (non-carcinogen)	15 (4)	0.2-0.48	0.34	0.13	0.34	1.13E+00	N/A	1.42E+12	7.79E-05	TR	N/A	TR
Fluoranthene (non-carcinogen)	17 (17)	0.012-3.09	0.80	0.87	0.49	3.94E+00	N/A	1.42E+12	7.79E-05	LN	N/A	LN
Xylene (non-carcinogen)	7 (10)	0.0074-68.7	10.14	20.95	4.77	-	4.55E+07	N/A	1.97E-02	N	N/A	N
Xylene (non-carcinogen)	8 (10)	0.053-7800	1952.43	3526.22	1.33	-	4.55E+07	N/A	1.97E-02	N	N/A	N
Xylene (non-carcinogen)	15 (3)	0.003-2	0.68	1.14	0.36	-	4.55E+07	N/A	1.97E-02	TR	N/A	TR
Xylene (non-carcinogen)	16 (8)	0.1-510	45.49	123.30	6.90	-	4.55E+07	N/A	1.97E-02	TR	N/A	TR
Xylene (non-carcinogen)	18 (15)	0.08-1.90	0.47	0.58	0.20	1.06E+00	4.55E+07	N/A	1.97E-02	LN	N/A	LN

(1) See Table 1 for a list of sites and site numbers.
(2) See Chapter 2 for an explanation of Lands method of calculating the UCL_{95}. (-) Indicates that the Maximum value was used for intake dose calculations.
(3) See Chapter 3 for an explanation of the volatilization factor (VF).
(4) See Chapter 3 for an explanation of the particulate emissions factor (PEF).
(5) See Chapter 3 for an explanation of the groundwater mass transfer parameter (GW).
(6) See Chapter 3 for an explanation of the @Risk best-fit function and methods for determining the best fit function. TR-triangular distribution, N-normal distribution, LN-lognormal distribution, U-uniform distribution.
(7) Indicates that the distribution function listed below are for contaminant concentrations in the soil.
(8) Indicates that the distribution function listed below are for contaminant concentrations in the air.
(9) Indicates that the distribution function listed below are for contaminant concentration in the groundwater.

Vita

Arne Edward Olsen

Arne Edward Olsen was born in Lynchburg, Virginia on November 10, 1973. He is the son of Arne and Rebecca Olsen of Concord, Virginia. Arne received his undergraduate degree from Virginia Polytechnic Institute and State University (Blacksburg, VA) in Environmental Science, May of 1996. After graduation Arne worked for Alliant Powder and Ammunition Co. LLC at the Radford Army Ammunition Plant as an environmental engineer. In August of 1999 Arne returned to Virginia Polytechnic Institute and State University to begin his Masters in Environment Science and Engineering working with Dr. Naraine Persaud.