12 UNCERTAINTY ANALYSIS AND RESULTS

12.1 Introduction

This dose reconstruction was conducted in two steps: 1) a deterministic step resulting in a point-estimate of dose for each receptor and 2) a probabilistic step resulting in many estimates of dose (a dose distribution) for each receptor. The step 1 dose estimates, presented in Chapter 11, Point-Estimate Results, were generated by specifying a single, nominal value for each variable in the models used to estimate dose. This resulted in a single dose estimate for each of the hypothetical receptors. The nominal values for each variable were considered to be realistic and representative of conditions at the SRS; however, some of the hypothetical scenarios postulated conditions and behaviors likely to bound actual doses. Thus the nominal dose estimates (point estimate results), from this first step of the study, are believed to be representative of doses for the hypothetical receptors. However, when the second step of the analysis considered uncertainty in the input variables, the mean and median values of dose that resulted generally exceeded the nominal values developed in step 1. This does not appear to be an anomaly, but can be attributed to the nonlinear behavior of the dose calculation when uncertainty is considered. Although the step 1 (point estimate) analysis provides important information on the radionuclides, pathways, and receptor behaviors determining the numerical results, the step 2 (uncertainty) analysis likely provides a more reliable estimate of dose.

While the single-point estimates used in a deterministic estimate of dose may be considered "representative" or "realistic," they do not capture the inherent uncertainty in the variables of interest. By nature, many of the natural processes and phenomena that investigators attempt to represent with equations and numbers will vary with time, space, and environmental conditions. Likewise, many of the variables used to represent or describe a natural process will change with time, space, and the environmental conditions.

In addition to the uncertainty in the natural phenomena, there is uncertainty associated with our understanding of how the phenomena used to represent natural processes are related (the conceptual models), their mathematical descriptions, and their computational implementation. When these uncertainties are considered, the reconstructed dose is uncertain; instead of a single-point estimate of dose for each receptor (as described in Chapter 11), multiple values (a probability distribution of dose) result. Simply, the uncertainty approach involves treating specified input variables as sets of values rather than as single values in the computation. The computation is repeated multiple times using various combinations of the input variable values. This results in a set of dose estimates rather than a single value. This set can then be used to describe the uncertainty associated with the result from the deterministic approach.

In the dose reconstruction for the Savannah River Site, the analysts focused on the uncertainty associated with the variables that are used to describe a particular process (e.g., the bio-uptake factors, consumption of foodstuffs, and particle size). This approach is the most common way to address uncertainty because it does not require development of alternative conceptual models or computational codes. In addition, this approach enables an analysis of the sensitivity (i.e., what changes in the results [dose] are produced by specific changes in the input variables). In this report, the term "uncertainty analysis" is used to describe three types of evaluation:

1. Description of the uncertainty in dose (i.e., how much variability is estimated for dose given uncertainty in the input variables).
2. Description of how the uncertainty in dose depends on the uncertainty for each uncertain input variable.

3. Description of the sensitivity of dose to variations in input variables.

Sometimes the terms "probabilistic" or "stochastic" analysis are used to describe these evaluations.

12.1.1 Overall Description of Uncertainty Analysis

Uncertainty analysis is sometimes referred to as probabilistic analysis or Monte Carlo analysis. It involves multiple computations that use the same code but change the values of the input variables with each computation. The input values are selected randomly from either sets of observations or distributions, using observations and judgment based upon experience. Section 12.1.1.2 describes this selection process. The selected values for the input variables of interest are then used in the computation, and the value for each of the sampled variables changes with each computation. For example, if 10 values are selected for each variable, the computation is repeated 10 times. It should be noted that each individual computation is a deterministic analysis similar to the point-estimate case but with different values for the variables.

In this analysis, the effect of 14 variables on dose was investigated. This required a series of computations that changed the values of these variables with each computation. As explained below, each computation is called a realization and yields a dose estimate based on a particular choice of values for all the input variables.

In an uncertainty analysis, the aggregate results are generally of interest rather than the result from any one computation. This set of results can be described in terms of simple statistics to give the representative values (e.g., mean), range, and nature of the distribution. In exercises such as dose reconstruction or risk analysis, uncertainty analysis provides insight into the uncertainty associated with the results.

Uncertainty analysis can be approached in a variety of ways. The most common approach is the Monte Carlo approach, which is based on random selection of the variable values from specified distributions. Section 12.1.1.2 describes the details of the Monte Carlo approach.

12.1.1.1 Rationale – Why We Do It; What Are We Trying to Determine?

The primary task of this project was to estimate the dose to 28 hypothetical individuals resulting from exposure to releases from the Savannah River Site (SRS). This estimation was performed using two methods: a point-estimate analysis and an uncertainty analysis. The point-estimate analysis used a representative value for each of the model input variables and resulted in an estimate of dose, cancer incidence risk, and cancer fatality risk for each of the 28 hypothetical receptors specified in the 7 scenarios. The point-estimate analysis included all credible exposure pathways given the behaviors specified in the scenarios and the land-use, water-use, and agricultural practices in the region surrounding the SRS during the period studied. The point-estimate analysis included only those radionuclides that survived the Level 1 screening of the Phase II report. Chapter 11 describes the point-estimate results. The point-estimate results provided a very detailed picture of what caused the dose and risk to each receptor including:

- How doses and risks changed in time.
- What radionuclides contributed most to dose and risks for a particular receptor.
- What pathways contributed most to dose and risks for a particular receptor.
- How the behaviors specified in the various scenarios influenced doses and risks.

An uncertainty analysis was undertaken with the objective of quantifying the uncertainty associated with the estimated doses. To reduce the computational and data-handling burden, risks were not carried along as part of the calculation; however, risks can be estimated from the doses using adult risk factors. Because

the uncertainty analysis involves repeating the dose assessments of the point-estimate analysis many times, a daunting amount of computation would be required if all input variables were considered in the uncertainty analysis. To reduce the effort associated with defining a distribution for each variable, the dose-assessment model used for the uncertainty analysis was simplified. Section 12.2 describes this simplification.

Because the results of the point-estimate analysis were used to guide these simplifications, the point-estimate analysis played the role of a screening analysis. As such, the point-estimate analysis was used to determine which pathways and radionuclides were very minor contributors to dose and could be excluded from the uncertainty analysis. In addition, the point-estimate results helped to determine which variables could be considered to be fixed, rather than uncertain, in the uncertainty analysis as described in Section 12.3.

12.1.1.2 Approach – Monte Carlo Sampling of Inputs to Provide Distributions of Dose

In uncertainty analysis, a randomly generated value selected from the probability distribution of each uncertain variable is assigned to that variable; other variables considered to be certain are assigned their nonrandom values. The values assigned to all variables comprise the set of values for one realization (i.e., one set of input variables for one computer run) that results in a random value as the outcome for that particular realization. Depending on the number of uncertain variables involved in the model and the sampling method adopted, the number of realizations required for good statistical estimates may range from a few to a few thousand. A distribution of dose would be obtained after estimating the dose for each of the realizations considered.

The process of random sampling falls into two categories: simple random sampling (Monte Carlo sampling [MCS]) and stratified sampling (e.g., Latin Hypercube Sampling [LHS]) (1). In MCS, each uncertain variable is assigned a range and probability distribution (mathematically described as a probability density function) that may be based on observations, judgment, or a combination of the two. The probability density functions for each of the uncertain variables of interest are then randomly sampled, resulting in a set of values for each of the input variables of interest. Generally, many thousands of samples are required to adequately represent the probability density function when using simple random sampling. A large number of samples are needed to obtain values over the entire range of the variable.

The LHS technique divides the probability density function associated with an uncertain variable into several strata ("bins") to ensure the probability of choosing a random value from each stratum is the same. Compared to the Monte Carlo sampling technique, LHS is more efficient and enables more complete sampling from the probability density function with a limited number of samples. This technique requires more computer memory because all samples are generated at once to ensure appropriate distribution and independence or correlation among the variables. In this analysis, the computation time for each realization was sufficiently long to make LHS the preferred sampling technique because fewer realizations are required.

12.2 Development of a Simplified Model

Uncertainty analyses can be very computationally intensive, requiring a number of individual computations (realizations) that is many times the number of input variables. Each realization requires another computation using the preprocessor, GENII v.2 Code, and postprocessor as described in Chapter 4. To make the problem more manageable, it was decided to focus the uncertainty analysis on those radionuclides and pathways that were most important in the determination of the point estimate. This simplification reduced the computation time and the amount of data handling.

All the various contributors to dose, as described in Chapter 11, were evaluated. These include specific pathways such as dose from eating vegetables or dose from a specific radionuclide such as Cs-137. Each pathway and/or radionuclide contribution was compared with the total dose for a particular receptor. The contributors to dose by pathway and radionuclide were arranged in order from least to greatest. The smallest contributor was eliminated, then the next smallest was eliminated, and the process was repeated until the cumulative contribution of eliminated pathways and/or radionuclides to the total dose was just under 5 percent. This process was performed on all 28 receptors.

The set of radionuclides and pathways satisfying the 5-percent criterion for all receptors was chosen as the set for simplifying the analysis. Because the pathways and radionuclides were not mutually exclusive (e.g., removing both soil ingestion and uranium takes out uranium dose from soil ingestion only once, but it was counted twice for the cumulative percent), the actual reduction in total dose was always less than 3 percent. With the removal of a particular radionuclide or pathway, a number of input variables were also eliminated from the computation. This helped to reduce the number of variables that needed to be evaluated for inclusion in the uncertainty analysis.

It should be noted that eliminating radionuclides and pathways from the analysis of a particular scenario in effect changed the conceptual model that was evaluated for the scenario. The effect of this change in the conceptual model was evaluated by comparing the point estimate of total dose (all pathways and radionuclides) to the estimate of total dose from the simplified model (reduced set of pathways and radionuclides). The total dose (using the simplified model but using the exact same input variable values as the point estimate) is designated "Run 0."

Table 12-1 presents the comparison of dose for each receptor from both the original point estimate and from the simplified computation ("Run 0"). Because the simplified computation provides doses within 3 percent of the more complete model, this simplification was considered to be acceptable (Table 12-1 shows that the largest deviation is 2.6 percent for the Child Born in 1964 for the Migrant Family scenario).

Table 12-1 Comparison of Dose Estimates from Complete and Simplified Models

Scenario	Family Member	Point Estimate (millisieverts)	Run 0 (millisieverts)	Ratio, Run 0 over Point Estimate
Delivery Family	Adult Female	6.106	6.091	0.997
Delivery Family	Adult Male	6.283	6.266	0.997
Delivery Family	Child Born 1955	5.200	5.180	0.996
Delivery Family	Child Born 1964	2.090	2.081	0.996
Migrant Family	Adult Female	0.447	0.438	0.979
Migrant Family	Adult Male	0.624	0.614	0.983
Migrant Family	Child Born 1955	2.178	2.160	0.992
Migrant Family	Child Born 1964	0.083	0.081	0.974
Near Water Family	Adult Female	2.091	2.057	0.984
Near Water Family	Adult Male	2.205	2.170	0.984
Near Water Family	Child Born 1955	3.137	3.099	0.988
Near Water Family	Child Born 1964	1.759	1.734	0.986

Scenario	Family Member	Point Estimate (millisieverts)	Run 0 (millisieverts)	Ratio, Run 0 over Point Estimate
Outdoor Family	Adult Female	3.030	3.001	0.990
Outdoor Family	Adult Male	4.216	4.169	0.989
Outdoor Family	Child Born 1955	9.435	9.383	0.994
Outdoor Family	Child Born 1964	1.826	1.810	0.991
Rural Family One	Adult Female	0.303	0.299	0.985
Rural Family One	Adult Male	0.423	0.418	0.987
Rural Family One	Child Born 1955	1.589	1.580	0.994
Rural Family One	Child Born 1964	0.072	0.071	0.981
Rural Family Two	Adult Female	0.696	0.685	0.985
Rural Family Two	Adult Male	0.974	0.961	0.987
Rural Family Two	Child Born 1955	3.751	3.729	0.994
Rural Family Two	Child Born 1964	0.140	0.137	0.979
Urban Family	Adult Female	0.330	0.325	0.986
Urban Family	Adult Male	0.731	0.723	0.989
Urban Family	Child Born 1955	2.686	2.675	0.996
Urban Family	Child Born 1964	0.107	0.106	0.983

12.3 Input Variables and Realizations

As discussed in the previous sections, the input variables involved in the computational model for the uncertainty analysis fall into two categories: 1) input variables that are considered as uncertain and 2) input variables that are considered as certain or fixed. While all input variables have an inherent uncertainty, the analysis described in Section 12.4 identified those variables that have the largest effect on the uncertainty of the resultant dose. This analysis reduced the number of input variables that will be considered as uncertain in the uncertainty analysis to 14, leaving the remaining input variables to be considered as certain (fixed) and thus treated as point estimates or fixed values.

Certain and uncertain input variables are introduced to the computational model differently. Certain input variables are represented by a single value, whereas an uncertain input variable is represented by a probability distribution and its associated statistics. Because an uncertain input variable cannot be represented by a single value, a set of values sampled from the variable's probability distribution is used to represent the variable of interest. The degree of representativeness depends on the sampling technique and the number of samples taken from the probability distribution.

As noted in Section 12.1.1.2, random sampling techniques fall into two categories: 1) simple random sampling (SRS), also known as MCS, and 2) the stratified sampling, also known as the LHS. The MCS technique does the sampling completely at random with each sample taken sequentially, whereas the LHS technique performs the sampling by a constraining value selection based on previously constructed realizations. By not allowing samples to be drawn from previously sampled intervals for a variable, LHS provides a more representative sampling of the distributions with a smaller sample size.

MCS is a computationally time-intensive sampling technique. The samples are chosen completely at random within the range of the probability distribution. This necessitates large numbers of random samples for highly skewed or long-tailed probability distributions to reasonably represent the uncertain nature of the variable under consideration.

LHS is more efficient than MCS because it is designed to accurately recreate the input distribution with fewer samples than the MCS method. However, LHS is a memory-intensive technique because the entire set of samples for all variables is collected at once.

All input values (fixed and sampled) are collected together in a matrix format. One set of input values from this collection, with a single value for each variable, is referred to as a realization. Each realization, when used in the computer code, results in a single value for the output variable (e.g., dose). When an uncertainty analysis is performed, the computer code is run iteratively, each time using a different combination of input variable values as determined by the sampling. For an uncertainty analysis involving n uncertain variables, 3n realizations (computer runs) are generally considered to be adequate when using LHS sampling.

12.3.1 Description

LHS was adopted as the sampling technique for the uncertainty analysis. After examining two general-purpose commercial software packages for risk analysis—Crystal Ball 2000 (2) and @RISK (3), it was decided to use the LHS computer code developed by Sandia National Laboratories for the U.S. Nuclear Regulatory Commission (1). This decision was based on the need to be able to define correlations among the uncertain variables. The LHS computer code requires each uncertain variable to be identified by its probability density function (probability distribution) and two quantiles for that distribution: the 0.001 quantile (0.1 percentile) and the 0.999 quantile (99.9 percentile).

Table 12-2 provides the 14 variables that were considered to be uncertain for the uncertainty analysis. For each variable, Table 12-2 contains the type of probability distribution and some of the statistics for that distribution, including the 0.001 and 0.999 quantiles. These quantiles were obtained from built-in functions of Microsoft Excel although these quantiles may lay outside the range of reality for a particular variable they were the required input for the LHS code

The 14 uncertain variables involved in the uncertainty analysis were considered to be independent of one another. Therefore, during the preparation of the input data file for the LHS code (discussed in the previous section), a particular “flag” was used to indicate that no correlation existed among the variables involved. This flag resulted in small correlation coefficients (less than 0.2) when the correlation coefficient matrix associated with the 40 X 14 output matrix (the output matrix with 40 realizations) was examined.

12.4 Overall Description of the Simplification Approach

This section describes the process by which 14 variables were selected for uncertainty/sensitivity analysis from an initial list of 331 possible variables.

12.4.1 Reasons for Limiting the Scope

A standard reference on radiological assessment (4) states:

> “The first step in an uncertainty analysis is to limit the scope. This requires an explicit statement of the objectives of the assessment and a determination of relevant radionuclides, exposure

> pathways, and model parameters. Limiting the scope of an uncertainty analysis avoids exhausting financial, physical, and human resources on aspects of assessment models that are not significant."

For this analysis, the scope was limited to focus attention on variables with the greatest potential for affecting variation in the dose to receptors. This avoided unnecessary use of resources for insignificant aspects of the models.

By limiting the analysis scope to those radionuclides, pathways, and input variables most influencing dose and dose variation, the computational burden was focused on the most important model aspects (some of this simplification is described in Section 12.2). By screening the input variables to select those influencing variations in dose most, the development of probability distributions was focused on the most important variables.

Table 12-2 The Fourteen Uncertain Variables Considered for Uncertainty Analysis

Variable	Description	Units	Dist. Type	Pt. Estimate	Log Mean	Log Std. Dev.	Median	0.001 Quantile	0.999 Quantile
Water Pathway									
CLBFF, Cs-137	Bioconcentration in fish	L/kg	Lognormal	4,700	8.46	1.20	4.70E+03	116	1.90E+05
CLBFF, Sr-90	Bioconcentration in fish	L/kg	Lognormal	450	6.11	2.03	4.50E+02	0.849	2.39E+05
CLBFF, P-32	Bioconcentration in fish	L/kg	Lognormal	50,000	10.82	0.89	5.00E+04	3.15E+03	7.93E+05
Air Pathway									
LEAFRS, RESFAC;1	Soil resuspension factor—farm	1/m	Lognormal	0.00001	-11.51	2.62	1.00E-05	3.03E-09	3.30E-02
RADIUS	Particle radius	µm	Lognormal	0.5	-0.69	0.41	5.00E-01	0.143	1.75
WTIM	Weathering rate constant from plant	d	Lognormal	14	2.64	1.40	1.40E+01	0.183	1.07E+03
DPVRES	Deposition velocity from air to plant surfaces	m/s	Lognormal	0.001	-6.91	1.26	1.00E-03	2.07E-05	4.84E-02
DEPFR1, 2	Deposition retention fraction for plants	Fraction	Lognormal	0.25	-1.39	0.59	2.50E-01	4.07E-02	1.54
GRWPA, 5&6	Growing period for animal forage	d	Lognormal	30	3.40	0.44	3.00E+01	7.65	118
CLFMT, I-131	Cattle intake-to-beef transfer factor	d/kg	Lognormal	0.04	-3.22	0.50	4.00E-02	8.49E-03	0.188
F	Release factor		Lognormal	1	0.00	0.20	1.00E+00	0.539	1.86
CONSUM, 5	Animal feed consumption rate— beef animal forage	kg/d	Normal	36	36.00*	8.63*		9.33	62.7
CLFMK, I-131	Cow intake-to-milk transfer factor	d/L	Lognormal	0.01	-4.61	0.91	1.00E-02	6.06E-04	0.165
BIOMA2, 5&6	Animal forage standing biomass (wet)	kg/m^2	Lognormal	0.3	-1.20	0.08	3.00E-01	0.233	0.386

* Mean and standard deviation for normal probability distribution.

12.4.2 Variable Selection Process—Overview

The point estimates provided in Chapter 11 are based on a single value chosen for each input variable. All input variables were considered to be fixed, or "certain," even though a range of values may have been considered in determining the single representative value for each variable. These representative values were chosen to be consistent with the specified scenarios, and site characteristics. When a range of values was available for a variable, the representative value was chosen in such a way that it did not intentionally underestimate or overestimate the dose.

In the uncertainty analysis, all of the "certain" input variables used to generate the point estimates could be considered as uncertain variables with their own probability distributions because, in fact, all input variables are uncertain. However, in most examples of environmental analysis, only a few input variables dominate the results. This variable selection process is intended to reduce the number of variables treated as uncertain to focus attention on those that appear to have the most effect on variations in dose.

Two categories of input variables were categorically excluded from consideration as uncertain:

1. The variable values representing the behavior of the receptors (e.g., the amount of milk consumed at a particular age) were derived from the scenario specifications. Because the scenarios were hypothetical and specified by the CDC and the SRSHES, and because the specified behaviors represented a range of plausible variation, these variable values were considered to be fixed for the purposes of this study.
2. This study has used a set of dose and risk factors established by national and international radiation protection organizations over many decades. These factors were treated as fixed because the uncertainty in these factors was considered small compared to that in other input variables and the uncertainty in these variables would be similar for any dose estimate.

Figure 12-1 shows the sequence of steps used by the variable selection approach to reduce the number of input variables considered in the uncertainty analysis.

At the beginning of the selection process, all input variables except those categorically excluded were considered to be candidates for treatment as uncertain variables. This stepwise process was intended to eliminate variables as candidates for the "uncertain" category based on a defined and defensible criterion at each step in the process. The following criteria were used at the six steps of the selection process:

1. Eliminate variables concerned with exposure pathways that are not used for modeling the SRS.
2. Eliminate variables that are only used for radionuclides and pathways that are minor contributors to dose.
3. Eliminate variables that are only used for radionuclides and pathways whose fractional contribution to dose squared is small.
4. Eliminate any remaining variables in categories already determined as fixed, such as variables associated with scenario specifications or dose and risk coefficients.
5. Combine input variables that may be specified separately but are more appropriately considered the same for this study.
6. Eliminate variables with small contributions to variance.

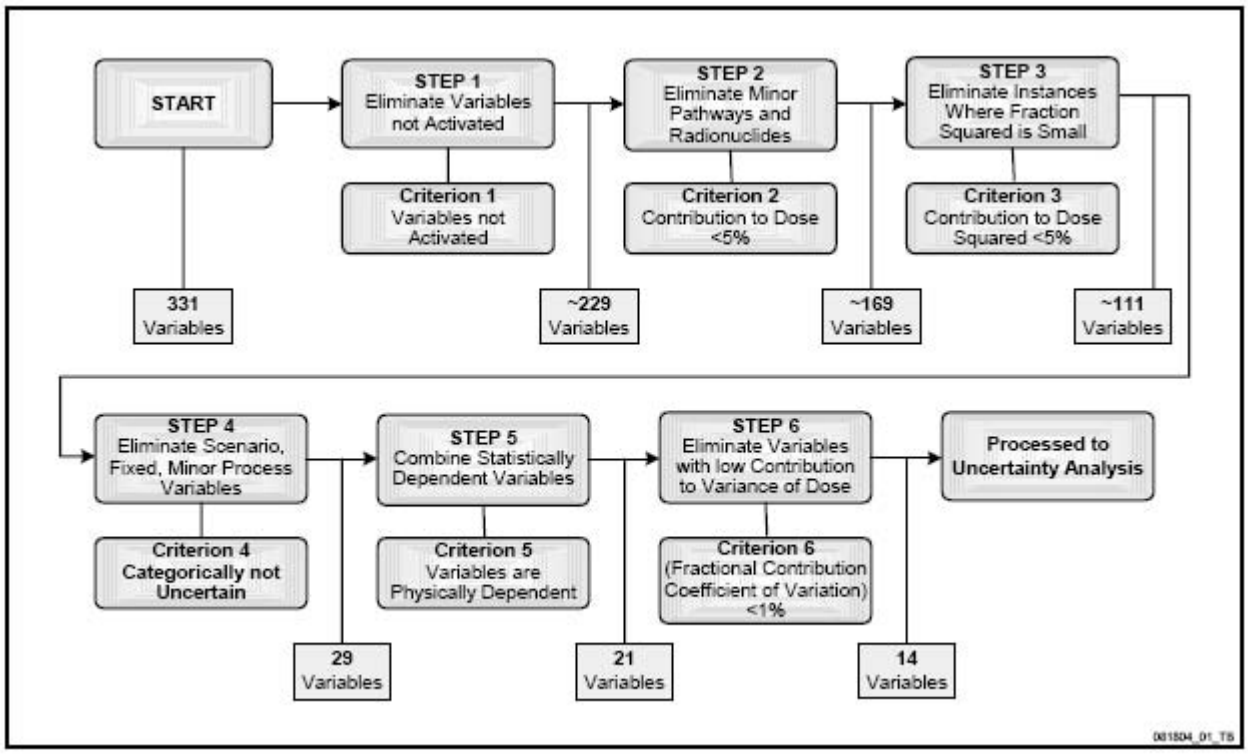


Figure 12-1 Steps Involved in Variable Reduction and Uncertainty Model Simplification

Each of these steps is described in more detail below. In accounting for the number of variables, Appendix F provides tables that list 331 separate variables. These lists are taken as the starting point of this process to reduce the number of variables considered as uncertain. However, there is a subjective element in counting the variables. For example, the variable "ARMETFILE" is the name of the file containing the 20-year average meteorological data that is read into GENII. These meteorological data were not counted among the input variables required to be specified because 1) they were computed separately and not included in the variable tables and 2) they were kept constant for all computations.

The main reason for performing the uncertainty analysis is to estimate the confidence interval for the estimated dose to each receptor. The confidence interval is a commonly used measure to describe the uncertainty in a variable. A key factor in determining the confidence interval is the variance of the dose (or equivalently the standard deviation of the dose, which is the square root of variance). Some of the screening steps are based on the fact that even though a variable may be important in determining the value of dose, it may have little influence on the variance of the dose.

12.4.2.1 Start

The 331 variables were specified to obtain the point estimate of dose for each receptor. Although each variable was carefully evaluated and specified, counting them is somewhat subjective. As stated in Section 12.4.2, some variables are just the name of a file containing extensive but fixed values—these variables were not included in the count. Other variables were indexed by exposure location or radionuclide and were specified for each location and radionuclide with potentially different values. In this description, they were counted as a single variable if a single value was used for every instance. Additional inputs to the dose modeling that were not counted as variables include 1) air releases by year, isotope, and source; 2) computer grid coordinates; and 3) water concentrations by year, isotope, and

location. In addition to the variable lists in Appendix F, a variable "F" was added to the lists which accounts for uncertainty in radionuclide releases.

12.4.2.2 Step 1—Eliminate Unused Variables

Because of the scenario specifications and the assumptions made for modeling, the point-estimate analysis did not use all the input variables employed by GENII to model the transport and exposure processes. For example, Appendix F shows that food chain transport variables associated with irrigation are not used due to the lack of evidence that the Savannah River had ever been used for irrigation in the region. Variables such as irrigation time (IRTIMA, IRTIMR, IRTIMT) and irrigation rate (RIRR, RIRRA, RIRRR) are not specified for the point-estimate analysis and need not be considered in the uncertainty analysis. As part of the scenario implementation, beef cattle were assumed to consume no contaminated feed (beef cattle consumed contaminated forage-grass); therefore, about 30 food chain variables were eliminated. Also note that cerium, technetium, niobium, zinc, cobalt, sulphur, phosphorus, yttrium (a decay product of niobium), and zirconium are only released to water, so their terrestrial uptake factors may be eliminated; this eliminates eight variables for each element. Approximately 102 variables were eliminated in this step, leaving approximately 229 potential uncertain variables.

12.4.2.3 Step 2—Screen Out Minor Contributors to Dose

As described in detail in Section 12.2, a simplified model was obtained by screening out radionuclides and pathways whose combined contribution to total dose was less than 5 percent. The actual reduction in dose was always less than 3 percent for any receptor. The following radionuclides and pathways were eliminated from the model on this basis:

1. Air Release:
   a. Radionuclides: americium-241, cesium-137, iodine-129, strontium-89, strontium-90, and all isotopes of uranium.
   b. Pathways: ground contamination, grain, and soil ingestion.
2. Water Release:
   a. Radionuclides: iodine-129, niobium-95, strontium-89, and all isotopes of uranium.
   b. Pathways: Boating, swimming immersion, and swimming inadvertent ingestion.

By eliminating these pathways and isotopes, it was unnecessary to consider about 60 associated variables as uncertain. For example, terrestrial uptake factors for various forms of vegetation and animals, and for the isotopes americium-241, cesium-137, iodine-129, strontium-89, strontium-90, and all uranium isotopes, could be eliminated from consideration as uncertain variables. Because uptake factors are generally element-specific and not isotope-specific, the uptake factors for iodine remained because iodine-131 was not screened out. Therefore, 8 terrestrial uptake factors for each of the elements americium, cesium, strontium, thorium (a decay product of uranium), and uranium were eliminated, thereby reducing the number of potential uncertain variables by 40. In addition, six variables related to grain and eight variables associated with direct exposure to contaminated soil were reclassified. Another six variables related to grain and aquatic uptake were eliminated. Approximately 60 variables were eliminated in this step, leaving approximately 169 potential uncertain variables.

12.4.2.4 Step 3—Screen Out Instances Where Fraction Squared Is Small

Appendix L discusses a mathematical approach in which the dose for any receptor is considered to be approximated by a linear function of each of the input variables, expanded about their nominal values

(i.e., the point-estimate values). Quite a bit is known about the point-estimate dose, including the fractional contribution to the total dose from each radionuclide and through each pathway. Mathematically, the fraction of dose F(ij) is known for each pathway (i) and each radionuclide (j). This information has been used in the variable selection process described in Section 12.2 to eliminate pathways and radionuclides from the model.

The approach discussed in Appendix L shows that the variance of the dose is proportional to the square of the sum of the fractional contributions to dose, where the sum is taken over all pathways and all radionuclides for which the variable has an impact on dose. Note that a particular variable may not have an impact on the dose resulting from a particular radionuclide or pathway. For example, the variable characterizing the uptake of cesium-137 by fish does not have any impact on the dose from iodine from drinking milk. As an approximation, it was assumed that if the sum of the squares of the fractions for all isotopes of a given pathway were small, then the pathway and its associated variables would not contribute significantly to the variance of dose. Similarly, it was assumed that if the sum of the squares of the fractions for all pathways of a given isotope were small, then the radionuclide and its associated variables would not contribute significantly to the variance of dose.

For air releases, four receptors were used as indicators to test for these conditions. The receptors were the Adult Female and Child Born in 1964 for Rural Family #2 and for the Migrant Worker Family. As discussed in Chapter 11, for air releases, the two adults and Child Born in 1955 had similar contributors to total dose; however, the Child Born in 1964 was different because that receptor missed the large iodine releases early in the SRS history. Consequently, fractional dose contributions to both the Adult Female and Child Born in 1964 were examined for this step of the variable reduction. Rural Family #2 was selected because it had fractional contributions to dose typical for most of the scenarios receiving doses from air releases only; however, the Migrant Family was somewhat of an outlier and was included for that reason. From these instances, certain radionuclides and pathways were found to be insignificant contributors to the variance of dose. The incremental contribution to total dose by every radionuclide-pathway pair was computed and then renormalized by the total dose from the pathways and radionuclides remaining after the model simplification described in Section 12.2. (As shown by the ratios in Table 12-1, this adjustment in fractional contribution was small when compared to the fractional contributions stated in Chapter 11.) These renormalized fractional values were then squared, and the squared fractional values were renormalized by the largest value. The renormalized squared fractions were then summed over pathways and radionuclides. In general, if any of the resulting sums were smaller than 0.01, then the variables associated with those radionuclides and pathways were eliminated. On this basis the variables associated with the following pathways or radionuclides were thus eliminated:

- Carbon-14.
- Plutonium-238,239.
- Ruthenium-106.
- Resuspended soil inhalation.
- Poultry ingestion.
- Eggs ingestion.
- Fruit ingestion.
- Root vegetable ingestion.

For the Child Born in 1964 for Rural Family #2, the sum for root vegetable ingestion slightly exceeded the 0.01 criterion.

Two receptors included in water exposure scenarios were used as indicators to test for these conditions. These were the Adult Female and Child Born in 1964 for the Delivery Family. The Delivery Family was chosen because it had the largest doses from water releases; the Adult Female and Child Born in 1964

were chosen for the reasons stated previously. From these instances in which water release doses dominated, the following pathways and radionuclides were found to be insignificant contributors (as stated above) to the variance of dose:

- Shoreline direct exposure.
- Cesium-144.
- Cobalt-60.
- Cesium-134.
- Hydrogen-3.
- Iodine-131.
- Plutonium-238,239.
- Ruthenium-106.
- Zinc-65.
- Zirconium-95.

By eliminating variables associated with these pathways and radionuclides that are insignificant contributors to the variance of dose, approximately 58 variables were reclassified as fixed. For example, 7 uptake factors each for the elements plutonium and ruthenium were eliminated, thereby reducing the number of potential uncertain variables by 14. Elimination of the egg and poultry pathways reduced the number of variables by about 22. Elimination of water-release radionuclides reduced the number of uncertain variables by nine. Elimination of variables associated with ingestion of fruit and root vegetables reduced the variable count by about 12. This left about 111 potential uncertain variables.

12.4.2.5 Step 4—Eliminate Health Effects Variables and Scenario-Based Variables

A number of the remaining potential uncertain variables were reclassified as variables to be considered certain or fixed for a variety of reasons. The following represent some of the important reasons for these categorical reassignments:

- Holdup times are unimportant for long-lived radionuclides; the dose from phosphorus-32 could be affected because it has a relatively short half-life (14.5 days), but it is a relatively small contributor.
- Holdup times for food delivery and processing were assigned as part of completion of the scenario specifications; because the scenarios are mandated, these variables are considered to be certain.
- It was independently determined that, for these scenarios, radionuclide uptake by vegetation from soil is small compared to radionuclide uptake by vegetation from direct deposition; therefore, all variables associated with modeling soil uptake are considered fixed. Approximately 48 variables were reclassified on this basis.
- Variables associated with household use of river water were considered to be fixed; although volatilization of radionuclides was considered for the Near Water Family, it accounted for a very small fraction of the total dose.

The variables that relate to these points were considered fixed and therefore were eliminated at this step. Approximately 82 variables were thus eliminated, thereby leaving approximately 29 potentially uncertain variables.

12.4.2.6 Step 5—Combine Variables that May be Considered Together

Further reduction in the number of variables considered to be uncertain was accomplished by noting that for three pairs of variables only one value, albeit an uncertain value, need be assigned to the pair:

1. BIOMA2, 5 and BIOMA2, 6.
2. GRWPA, 5 and GRWPA, 6.
3. DEPFR1 and DEPFR2.

For the first two pairs, the variables referred to values for milk and beef animal forage. Because the same grass was used for both milk and beef animal forage, these variables could be assumed to be equal. For the third pair of variables, dry- and wet-deposition fractions were assumed to be equal.

In the following two cases, the value of one variable in a pair was derived from the sampled value of the other variable in the pair by multiplying the first variable by a constant; essentially, the pair of variables is assumed to be different by a multiplicative constant and perfectly correlated as explained in the following:

1. CONSUM,5—the consumption rate of forage by beef cattle was assumed to be the fraction 36/29 the value of CONSUM,6—the consumption rate of forage by milk cattle. This is the ratio of the point estimate values. The assumption means that whatever causes the beef cattle to eat more, it will have the same effect on dairy cows.

2. RESFAC,1—the resuspension factor for rural locations is assumed to be 100 times the value of RESFAC,2—the resuspension factor for urban locations. It is essentially assumed that whatever forces produce more resuspension in the city and suburbs (e.g., higher wind) will do the same on the farm. For the point-estimate case, it was assumed that the resuspension factor in the rural areas was 100 times that in the city due to plowing and other agricultural activities. LEAFRS is essentially the same variable as RESFAC, so those variables were set equal to the corresponding values for RESFAC.

Both particle radius and density are factors in determining deposition of radionuclides, but the equation uses radius squared. Because radius should have a more significant effect if varied, it was chosen as the uncertain variable, and density was considered fixed.

The net effect of these various consolidations was to eliminate 8 variables and leave 21 independent candidates for variables to be treated as uncertain. These variables are listed in Table 12-3.

Table 12-3 Characteristics of Input Variable Distributions

	Variable	Description	Units	Type	p1*	p2*
Water Pathway						
1	CLBFF, Cs-137	Bioconcentration in fish	L/kg	Lognormal	8.4553	1.1975
2	CLBFF, Sr-90	Bioconcentration in fish	L/kg	Lognormal	6.1092	2.0300
3	CLBFF, P-32	Bioconcentration in fish	L/kg	Lognormal	10.8198	0.8945
Air Pathway						
4	CLFMK, I-131	Cow intake-to-milk transfer factor	d/L	Lognormal	-4.6051702	0.9069765
5	CLFMT, I-131	Cattle intake-to-beef transfer factor	d/kg	Lognormal	-3.2188758	0.5015594
6	ABSHUM	Absolute Humidity	kg/m^3	Normal	0.01125	0.00053
7	BIOMA2, 5&6	Animal feed standing biomass (wet)-beef & milk animal forage	kg/m^2	Lognormal	-1.20397	0.0813929
8	BIOMA2, 3	Animal feed standing biomass (wet)-Milk animal feed	kg/m^2	Lognormal	1.43156†	0.02909†
9	BIOMAS, 1	Standing biomass (wet) - Leafy vegetables	kg/m^2	Lognormal	0.03486‡	
10	WTIM	Weathering rate constant from plant	d	Lognormal	2.63906	1.40311
11	CONSUM, 5&6§	Animal feed consumption rate-Beef & milk animal forage	kg/d	Normal	36.0000	8.6300
12	CONSUM, 3	Animal feed consumption rate-Milk animal feed	kg/d	Lognormal	1.71†	0.262†
13	DRYFAC, 1	Dry/wet ratio - Leafy vegetables	fraction	Lognormal	0.10875†	0.00218†
14	GRWP, 1	Growing period - Leafy vegetables	d	Lognormal	0.16861‡	
15	GRWPA, 5&6	Growing period for animal feed - Beef & milk animal forage	d	Lognormal	3.40119738	0.4423365
16	GRWPA, 3	Growing period for animal feed - Milk animal feed	d	Lognormal	0.05103‡	
17	DEPFR1,2	Dry & wet deposition retention fraction to plants	fraction	Lognormal	-1.3862944	0.5873942
18	LEAFRS, RESFAC;1§	Resuspension factor from soil to plant surfaces - farm	1/m	Lognormal	-11.512925	2.6214129
19	DPVRES	Deposition velocity from soil to plant surfaces	m/s	Lognormal	-6.9077553	1.2555349
20	RADIUS	Particle Radius	µm	Lognormal	-0.6931	0.4055
21	F	Release Factor		Lognormal	0	0.2

§ Related by a constant factor.

* p1 and p2 are the log mean and log standard deviation of their corresponding lognormal distributions, respectively, except for variables No. 6 and 11 that they are the arithmetic mean and the arithmetic standard deviation of the normal distribution, respectively.

†Arithmatic mean and arithmetic standard deviation.

‡Ratio of arithmetic standard deviation over arithmetic mean.

12.4.2.7 Step 6 – Retain Only Variables Contributing Substantially to Variance

The final step in reducing the number of input variables to be used in the uncertainty analysis was based on a factor (Column D in Table 12-4) that is the squared product of 1) the coefficient of variation for the variable (Column A in Table 12-4 and 2) the fractional contribution to dose from all pathways and radionuclides (Column B in Table 12-4) for which the variable is involved.

The coefficient of variation is equal to the standard deviation of the distribution describing the variable divided by the arithmetic mean of the distribution. The coefficient of variation is a relative measure of the variability of a random variable with respect to its mean value. Because the coefficient of variation (CV) measures relative variability, it allows comparison of variability across variables with different units and different absolute magnitudes. A variable with a small CV is very repeatable and reliable, experiencing little relative change from one measurement to another. A variable with a large CV is much more uncertain.

The data used to determine probability distributions for the 21 variables remaining at this step were gathered from a variety of published sources (with the exception of one case for which a personal communication was used). Appendix M provides details of the development of these probability distributions.

Table 12-4 shows the 21 variables, the coefficient of variation, the fractional contribution, the product of these squared, and the rank of the variable based on the combined factor. The following observations are based on this table:

- For the water-release variables, the range in the determining factor (Column D) is less than two orders of magnitude between the largest and smallest value; therefore, all three variables were retained.
- For the air-release variables, the highest ranked variable was the resuspension factor (the determining factor is 664) because the fractional contribution was relatively large and the coefficient of variation was over 31. This is an extremely large value that ranges over several orders of magnitude.
- Consider eliminating any variable with a determining factor less than 0.01: For the air release, the determining factor value falls off rapidly, and if the criterion of less than 0.01 was used, all variables of rank less than 10 would be excluded.
- The cutoff for rank was chosen to be 10. However, the variable of rank 11 was also included because it participated in dose from milk and meat as well as the companion variables—growing period (GRWPA,5) and forage consumption (CONSUM,5).

Based on these considerations, a final number of 14 variables were considered to be uncertain. Table 12-2 provides the statistical characteristics of these 14 variables, and Section 12.3 discusses how they were used.

Appendix N provides the input data used for the computations (LHS and GENII computer codes).

Table 12-4 Final Screening Step

		Variable	Coefficient of Variation (Standard Deviation /Mean)	Fractional Contribution to Total for Medium	Approximate Contribution to Variance	Square of Contribution to Variance	Rank for Medium
Water Pathway			***A***	***B***	***C***	***D***	
✓	1	CLBFF, Cs-137	1.788E+00	5.25E-01	9.38E-01	8.79E-01	2
✓	2	CLBFF, Sr-90	7.786E+00	1.59E-01	1.24E+00	1.53E+00	1
✓	3	CLBFF, P-32	1.107E+00	2.68E-01	2.96E-01	8.79E-02	3
Air Pathway							
✓	4	CLFMK, I-131	1.130E+00	1.38E-01	1.56E-01	2.43E-02	10
✓	5	CLFMT, I-131	5.348E-01	5.09E-01	2.72E-01	7.41E-02	7
	6	ABSHUM	4.711E-02	6.05E-01	2.85E-02	8.12E-04	13
✓	7	BIOMA2, 5&6	8.153E-02	6.93E-01	5.65E-02	3.19E-03	11
	8	BIOMA2, 3	2.032E-02	3.45E-01	7.01E-03	4.91E-05	16
	9	BIOMAS, 1	3.486E-02	6.09E-02	2.12E-03	4.50E-06	17
✓	10	WTIM	2.482E+00	8.16E-01	2.03E+00	4.10E+00	3
✓	11	CONSUM, 5	2.397E-01	6.93E-01	1.66E-01	2.76E-02	9
	12	CONSUM, 3	1.532E-01	3.45E-01	5.28E-02	2.79E-03	12
	13	DRYFAC, 1	2.000E-02	6.09E-02	1.22E-03	1.48E-06	18
	14	GRWP, 1	1.686E-01	6.09E-02	1.03E-02	1.05E-04	15
	15	GRWPA, 5&6	4.649E-01	6.93E-01	3.22E-01	1.04E-01	6
	16	GRWPA, 3	5.103E-02	3.45E-01	1.76E-02	3.10E-04	14
✓	17	DEPFR1,2	6.419E-01	8.16E-01	5.24E-01	2.74E-01	5
✓	18	LEAFRS, RESFAC;1	3.104E+01	8.30E-01	2.58E+01	6.64E+02	1
✓	19	DPVRES	1.959E+00	8.16E-01	1.60E+00	2.55E+00	4
✓	20	RADIUS	2.913E+00	1.00E+00	2.91E+00	8.49E+00	2
✓	21	F	2.000E-01	1.00E+00	2.00E-01	4.00E-02	8

✓ = retain for final analysis

12.5 Results of Uncertainty Analysis

The computer runs provided 40 dose values (corresponding to 40 realizations) for each receptor, as provided in Appendix O. In other words, the dose for each receptor was computed 40 times as a function of 40 (random) realizations of the variables considered to be uncertain plus all the other input variables considered to be fixed. Each set of 40 output values was examined statistically to determine an empirical

probability distribution and values representative of the central tendency (e.g., mean and median) for the dose to each receptor.

12.5.1 Description of Dose Distributions

The lognormal probability distribution was assumed to be a potentially good fit to the 40 random dose values associated with each receptor. This was assumed because all except 1 of the 14 probability distributions for the 14 uncertain input variables considered were lognormal (see Table 12-2). Also, the point-estimate value for each of these 14 variables from the deterministic analysis had been set equal to the median of the distribution for the corresponding variable in the uncertainty analysis. For this reason, the medians of the 28 distributions (1 for each receptor) were expected to be "close" to those point estimates found in the deterministic analysis contained in Chapter 11. In addition, the following is a well-known property of lognormal distributions: If a random variable is defined as the product of two or more independent random variables and each of these is described by a lognormal distribution, then the product will also be described by a lognormal distribution.

Using the computer software "Crystal Ball" (2), the best fit to the probability distributions for the total dose to 14 out of the 28 receptors was determined to be lognormal. The lognormal distribution was the second or third best fit to the dose distributions for the remaining receptors. However, even when the lognormal distribution was not the first choice, the best-fit distributions were similar to the lognormal. The fact that the dose distributions for all the receptors are not lognormal is attributed to the fact that the dose is computed by a complex mathematical model of the uncertain input variables involved in the uncertainty analysis. Except for very simple functions of random variables (e.g., multiplication by a constant), the function of several random variables is seldom characterized by the same probability distribution as the underlying input variables, even if they are all the same. For example, for the Adult Female member of the Delivery Family, the best fit to the distribution of dose from only air releases and the best fit to the distribution of dose from only water releases are gamma distributions; however, the best fit to the total dose from all releases, which is the sum of these air and water doses, is a lognormal distribution.

Table 12-5 displays descriptive statistics for the sample of 40 total dose realizations for each of the 28 receptors. The table lists the mean, median, minimum, maximum, and the standard deviation of total dose for each receptor. Note that the mean dose is larger than the median dose for every receptor. Because we have a sample of 40 values of dose, there are 20 values smaller than the median and 20 values larger than the median. If we had a symmetric distribution of dose, the smaller values and larger values would extend about equally on either side of the median value. In our case, however, the lower values are limited by zero because dose cannot be negative and the higher values extend out to relatively high values. For example, for the Child Born in 1955 for the Urban Family, the difference between the maximum value of dose (30.8 millisieverts) and the median dose (2.55 millisieverts) is 28.3 millisieverts; however, the difference between the median dose (2.55 millisieverts) and the minimum dose (0.345 millisieverts) is only 2.20 millisieverts. Figure 12-2 and Figure 12-3 show this skewing toward higher values, depicting as examples the dose distributions for the Adult Female in the Rural Family 2 and the Delivery Family, respectively. The distributions are displayed both as histograms and as a fitted lognormal distribution. Because the distribution is not symmetrical and is skewed toward higher values, the mean dose is larger than the median dose.

The range of the distribution of dose depends on the choice of both the scenario and family member. For example, for the Child Born in 1955 for the Urban Family, the ratio of maximum dose to minimum dose is over 89; for the Child Born in 1964 for the same family, the ratio of maximum dose to minimum dose is only about 4. Similarly, the ratio by which the mean dose exceeds the median dose depends on both the scenario and family member. The minimum value of this ratio is 1.07 for the Child Born in 1964 for the

Urban Family, and the maximum value of this ratio is 1.69 for the Child Born in 1955 for the same family.

Table 12-5 Statistics on Total Effective Dose Equivalent for Different Receptors*

Family	Family Member	Mean (mSv)‡	Median (mSv)‡	Minimum† (mSv)‡	Maximum† (mSv)‡	Standard Deviation (mSv)
Delivery Family	Adult Female	12.812	9.072	1.813	60.911	12.23
Delivery Family	Adult Male	13.001	9.281	1.737	61.183	12.26
Delivery Family	Child Born 1955	10.164	7.993	1.767	35.010	7.76
Delivery Family	Child Born 1964	4.498	3.159	0.712	15.146	3.73
Migrant Family	Adult Female	0.793	0.562	0.138	4.615	0.90
Migrant Family	Adult Male	1.117	0.756	0.184	6.770	1.29
Migrant Family	Child Born 1955	3.676	2.489	0.417	24.269	4.45
Migrant Family	Child Born 1964	0.127	0.093	0.043	0.732	0.13
Near Water Family	Adult Female	3.431	2.738	1.138	9.896	2.04
Near Water Family	Adult Male	3.574	2.929	1.183	10.084	2.06
Near Water Family	Child Born 1955	4.815	4.311	1.293	18.333	3.03
Near Water Family	Child Born 1964	2.850	2.290	0.867	14.090	2.31
Outdoor Family	Adult Female	4.687	4.263	1.272	11.751	2.46
Outdoor Family	Adult Male	6.055	5.546	2.026	14.246	2.95
Outdoor Family	Child Born 1955	13.331	10.988	2.529	60.270	10.21
Outdoor Family	Child Born 1964	2.951	2.309	0.893	14.158	2.31
Rural Family One	Adult Female	0.502	0.387	0.090	3.005	0.53
Rural Family One	Adult Male	0.712	0.538	0.120	4.406	0.76
Rural Family One	Child Born 1955	2.681	1.697	0.281	17.410	3.34
Rural Family One	Child Born 1964	0.093	0.077	0.037	0.340	0.06
Rural Family Two	Adult Female	1.174	0.890	0.199	7.162	1.28
Rural Family Two	Adult Male	1.655	1.198	0.267	10.502	1.85
Rural Family Two	Child Born 1955	6.362	4.006	0.642	41.579	0.798
Rural Family Two	Child Born 1964	0.190	0.153	0.072	0.796	0.14
Urban Family	Adult Female	0.447	0.284	0.083	2.204	0.39
Urban Family	Adult Male	0.895	0.698	0.263	3.276	0.55
Urban Family	Child Born 1955	4.314	2.551	0.345	30.820	5.57
Urban Family	Child Born 1964	0.115	0.107	0.054	0.215	0.03

* The number of decimal places for values in the table are to allow easy display; the values should be considered to have a precision no greater than two significant digits.
† These minimum and maximum values are for this sample; another set of realizations will likely have different values. However, any sample minimum value is greater than or equal to the population minimum value, while any sample maximum value is less than or equal to the population maximum value.
‡ mSv = millisieverts.

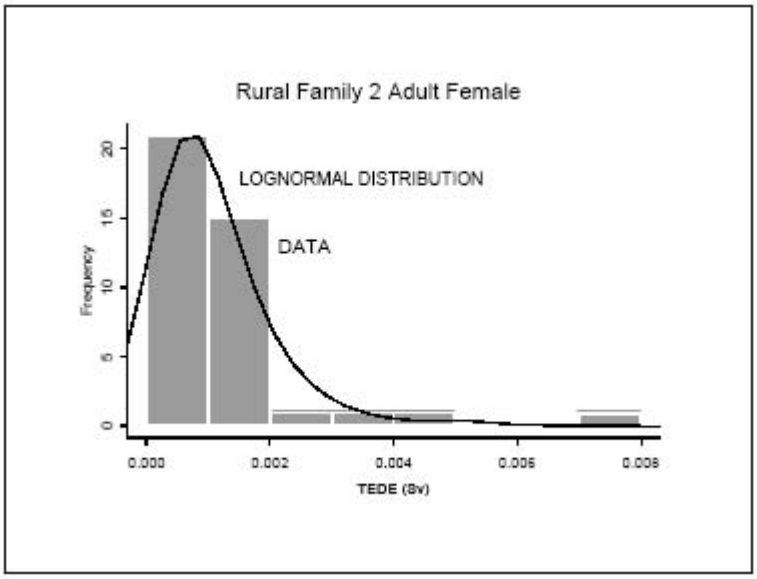


Figure 12-2 Histogram and Fitted Distribution for Dose to Adult Female, Rural Family 2

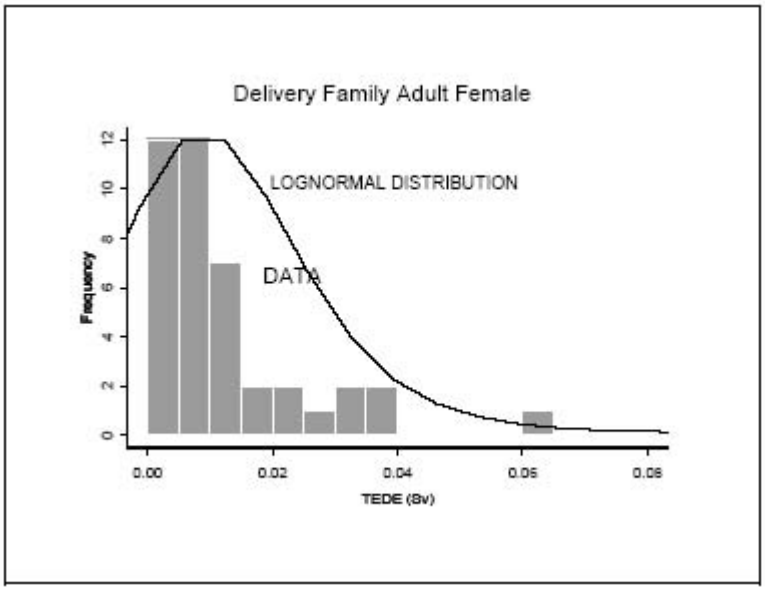


Figure 12-3 Histogram and Fitted Distribution for Dose to the Adult Female, Delivery Family

12.5.2 Quantified Confidence Bands for Doses

A primary motivation for the uncertainty analysis is to quantify the uncertainty associated with the dose to each receptor. One way to characterize this uncertainty is to state the confidence limits around the mean dose. Confidence limits define an interval around a parameter (e.g., the mean) so that the parameter is expected to be within the interval to a specified probability. Another way of looking at this is that the mean is estimated based on a finite set of values randomly sampled from a continuous distribution (i.e., in our case, we have a sample of 40 doses for each receptor out of an infinite number of possible values of

dose). The estimated parameter has two sources of uncertainty: 1) the underlying variability of the distributed values and 2) the finite sample size.

A standard statistical text (5) states the following:

$$\overline{X} - t_{1-\alpha/2}(n-1)\frac{S}{\sqrt{n}} \le \mu \le \overline{X} + t_{1-\alpha/2}(n-1)\frac{S}{\sqrt{n}} \qquad (12\text{-}1)$$

where, μ is the population mean
$\overline{X}$ is the sample mean
$t_{1-\alpha/2}(n-1)$ is the Student's t-distribution with (n-1) degrees of freedom for a 100(1-α)% total confidence level [100(α/2)% and 100(1-α/2)% confidence limits on the left and right sides of the distribution, respectively]
S is the unbiased estimator of the standard deviation
n is the number of observations (sample size).

If the sample variance is computed by the formula:

$$\hat{\sigma}^2 = \sum_{i=1}^{n} \frac{(X_i - \overline{X})^2}{n} \qquad (12\text{-}2)$$

then the unbiased estimator for the population variance is given by:

$$S^2 = \frac{n}{n-1}\hat{\sigma}^2 \qquad (12\text{-}3)$$

and the standard deviation, of either the sample or population, may be found by taking the square root of the corresponding variance.

For our computations, n=40. Then the fraction in equation (12-3) is 40/39 = 1.0256; if we take the square root to find S, then the fraction becomes $1.0256^{1/2}$= 1.0127. This relates the unbiased estimator to the sample standard deviation. This is a small correction, but it is made due to its ease of accomplishment.

Table 12-6 shows the values for t distribution for some typical confidence levels (1-α) and the 40 realizations.

Table 12-6 Values for t Distribution for Some Typical Confidence Intervals

α	Lower Bound	Upper Bound	Confidence Level	t(40-1)
0.1	5%	95%	90%	1.684875315
0.05	2.50%	97.50%	95%	2.022688932
0.02	1%	99%	98%	2.425840648

Since n=40, $40^{1/2}$ = 6.32455532. Then for α = 0.05,

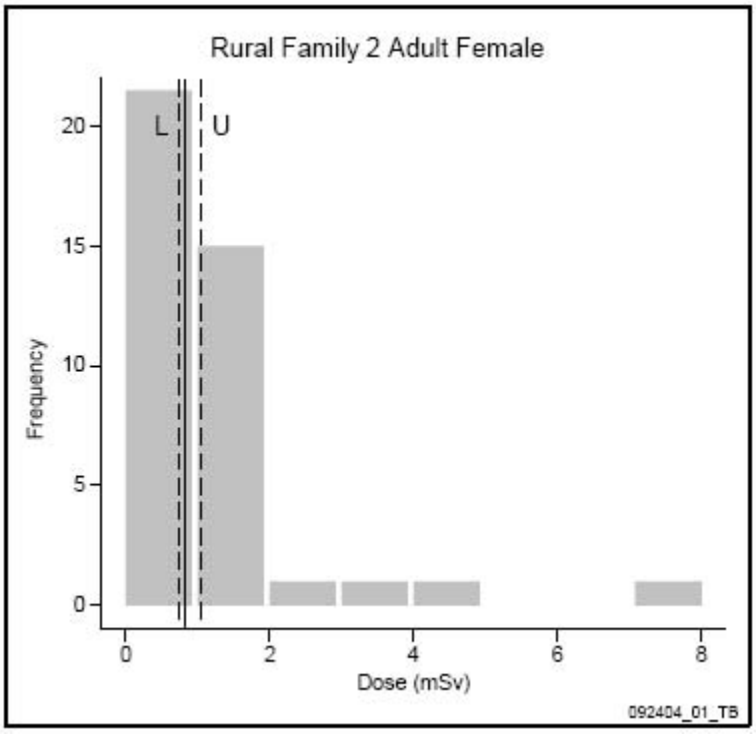
$$\overline{X} - 0.319815 \cdot S \le \mu \le \overline{X} + 0.319815 \cdot S \qquad (12\text{-}4)$$

Or the actual mean of dose is between the sample mean plus and minus about 1/3 of the sample standard deviation with a 95-percent confidence level. For example, for the Adult Female of the Delivery Family,

the sample mean dose is 0.012812 sievert; the actual mean dose is expected to be between 0.008899 sievert and 0.016724 sievert with a confidence level of 95 percent.

The median dose is closer to the point-estimate dose than to the mean dose because the medians of the lognormal distributions of input variables were set equal to the point-estimate values for those input variables. One way to make the same type of confidence statements about the median dose that have been made about the mean dose is to assume that the total dose distribution for each receptor is described by a lognormal distribution. A special property of the lognormal distribution is that the mean of the natural logarithm of the sampled doses is equal to the natural logarithm of the median dose. Therefore, equation 12-4 may be used to establish confidence limits around the ln (ln is used to represent the natural logarithm) median; by taking the inverse logarithm, one obtains the median and its confidence limits (i.e., raising e (the base of the natural logarithms) to the ln median and to the ln confidence limits provides the median and its associated confidence limits).

Table 12-7 displays the mean and median for all 28 receptors and also gives the upper and lower confidence bounds for each statistic at the 95-percent level. Figure 12-4 shows the confidence intervals around the median for one receptor (the Adult Female for Rural Family 2) overlaid on the histogram of dose.



U= Upper Bound
L= Lower Bound

Figure 12-4 Confidence Intervals Overlaid on Histogram of Dose

For the mean values, the confidence bounds are almost symmetrical; the confidence intervals on either side of the mean range from 9- to 41-percent of the mean value depending upon the scenario and family member. For the median values, assuming a lognormal distribution, the confidence bounds are skewed toward higher values (i.e., the difference between the upper bound and the median is generally a higher fraction of the median than the difference between the median and the lower bound). The lower bound of the confidence interval ranges from 9- to 26-percent of the median value; the upper bound of the confidence interval ranges from 9- to 36-percent of the median value.

Table 12-7 95% Confidence Intervals for Mean and Median of Total Effective Dose Equivalent for Each Receptor

Family	Family Member	Mean (mSv)‡	Lower Limit of 95% CI* for Mean (mSv)‡	Upper Limit of 95% CI for Mean (mSv)‡	Median † (mSv)‡	Lower Limit of 95% CI for Median (mSv)‡	Upper Limit of 95% CI for Median (mSv)‡
Delivery Family	Adult Female	12.812	8.899	16.724	8.662	6.464	11.608
Delivery Family	Adult Male	13.001	9.081	16.921	8.872	6.643	11.849
Delivery Family	Child Born 1955	10.164	7.681	12.646	7.826	6.173	9.923
Delivery Family	Child Born 1964	4.498	3.305	5.691	3.290	2.541	4.259
Migrant Family	Adult Female	0.793	0.504	1.082	0.554	0.431	0.713
Migrant Family	Adult Male	1.117	0.705	1.529	0.770	0.595	0.997
Migrant Family	Child Born 1955	3.676	2.252	5.100	2.465	1.871	3.246
Migrant Family	Child Born 1964	0.127	0.086	0.167	0.102	0.085	0.122
Near Water Family	Adult Female	3.431	2.778	4.083	2.953	2.480	3.517
Near Water Family	Adult Male	3.574	2.916	4.232	3.103	2.618	3.677
Near Water Family	Child Born 1955	4.815	3.846	5.784	4.150	3.493	4.932
Near Water Family	Child Born 1964	2.850	2.113	3.588	2.326	1.913	2.826
Outdoor Family	Adult Female	4.687	3.900	5.475	4.129	3.505	4.865
Outdoor Family	Adult Male	6.055	5.113	6.997	5.434	4.676	6.316
Outdoor Family	Child Born 1955	13.331	10.065	16.596	10.688	8.640	13.222
Outdoor Family	Child Born 1964	2.951	2.214	3.688	2.437	2.017	2.945
Rural Family One	Adult Female	0.502	0.332	0.672	0.367	0.289	0.465
Rural Family One	Adult Male	0.712	0.470	0.954	0.521	0.410	0.662
Rural Family One	Child Born 1955	2.681	1.612	3.750	1.765	1.333	2.337
Rural Family One	Child Born 1964	0.093	0.074	0.113	0.083	0.072	0.096
Rural Family Two	Adult Female	1.174	0.767	1.582	0.841	0.658	1.076

Family	Family Member	Mean (mSv)‡	Lower Limit of 95% CI* for Mean (mSv)‡	Upper Limit of 95% CI for Mean (mSv)‡	Median † (mSv)‡	Lower Limit of 95% CI for Median (mSv)‡	Upper Limit of 95% CI for Median (mSv)‡
Rural Family Two	Adult Male	1.655	1.064	2.246	1.166	0.905	1.501
Rural Family Two	Child Born 1955	6.362	3.808	8.915	4.150	3.121	5.519
Rural Family Two	Child Born 1964	0.190	0.144	0.236	0.164	0.141	0.192
Urban Family	Adult Female	0.447	0.322	0.572	0.349	0.281	0.433
Urban Family	Adult Male	0.895	0.718	1.073	0.784	0.670	0.918
Urban Family	Child Born 1955	4.314	2.532	6.096	2.659	1.955	3.616
Urban Family	Child Born 1964	0.115	0.104	0.125	0.111	0.101	0.121

* CI = confidence interval.
† Calculated based on the natural log values of total effective dose equivalent.
‡ mSv = millisievert.

Although the confidence intervals may be a substantial fraction of the central tendency value (the mean or the median), the upper confidence limits are not larger than the central tendency value (i.e., the confidence interval is always smaller than the interval 0 to twice the mean). This may seem counterintuitive to many who consider the dose estimates of this kind to be accurate only to an order of magnitude, if that. However, these confidence intervals reflect only the uncertainty in the input variables to the models and do not reflect the uncertainty in the models themselves or the approximations made in applying the models. In a sense, it is more correct to characterize these confidence bounds as a measure of precision rather than as a measure of accuracy.

For the dose distributions, the largest value of the coefficient of variation is approximately 1.3. This value is much smaller than the largest value of the coefficient of variation for the input variables, which was approximately 31 for the resuspension factor. The reason the uncertainty in resuspension factor is attenuated when processed by the dose models is that resuspension does not cause much dose for these scenarios. Inhalation of resuspended contamination has already been shown to be a minor pathway (see Section 12.2). In addition, resuspended contamination can cause dose by depositing on growing vegetation and becoming incorporated into the food chain. For the scenarios analyzed, however, direct deposition of contamination is always much more effective in causing dose than resuspended deposition.

12.5.3 Comparison of Results from Point Estimate and Uncertainty Analyses

Table 12-8 and Table 12-9 show for each receptor from the deterministic analysis the point-estimate dose compared to the results of the uncertainty analysis—the run 0 dose, the mean dose, and the median dose. As discussed previously in Section 12.2, the point-estimate dose is very close to the run 0 dose for each receptor -- as would be expected because the same values for the input variables are used. However, the mean dose from the uncertainty analysis is generally larger (and in some cases, substantially so) than the point-estimate dose for each receptor. The ratio of the mean dose to the point-estimate dose ranges from a high value of 2.15 (115 percent higher) for the Delivery Family Child Born in 1964 to a low value of 1.07 (7 percent higher) for the Urban Family Child Born in 1964. The median dose from the uncertainty analysis is generally higher than the point-estimate dose, but usually by a smaller amount than the corresponding mean dose. However, for the Urban Family receptors, the median doses are smaller than the point-estimate dose. The ratio of the median dose to the point-estimate dose ranges from a high value of 1.54 (54 percent higher) for the Delivery Family Child Born in 1955 to a low value of 0.86 (14 percent smaller) for the Urban Family Adult Female.

As discussed in Section 12.4.1, for the dose distributions obtained from the uncertainty analysis, it is not surprising that the median doses are generally smaller than the mean doses. However, it was somewhat surprising that the median doses were mostly larger (sometimes substantially so) than the point-estimate doses. After all, the point estimates of dose were developed using “representative” values for all the input variables, and a selected set of input variables were represented by distributions. All but one of the input distributions were chosen to be lognormal distributions (the other was a normal distribution). The median of each lognormal distribution was set equal to the point-estimate value of the corresponding input variable. Then the distribution of doses was generated based on 1) the distributions of input variables for those chosen to be uncertain and 2) the point-estimate values for those input variables chosen to be fixed. Because the medians of the input distributions were the point-estimate values, it could be expected (in a very approximate fashion) that the medians of the uncertain doses would be closer to the point-estimate values than the means of the uncertain doses. In some cases, however, they are greater than the point estimates by over 50 percent.

To examine this result, some additional analysis was performed. The dose to the Adult Female for the Delivery Person Family has been a focus of consideration because the water release dose (water dose) for

the point estimates is known to be 94 percent of the total dose. This receptor was chosen for closer scrutiny because it has dose dominated by this single release mode and because the median of the uncertain dose—9.072 millisieverts—is 1.486 times the point-estimate dose of 6.106 millisieverts.

An initial conjecture is that the air doses add to the water doses in a random fashion and provide a "floor" for the total dose. This more-or-less average addition could cause the low values of the total doses to be elevated and therefore increase the median. This conjecture does not seem to be confirmed by the data. The median for just the water doses is 8.539 millisieverts. This implies that only 5.9 percent of the median dose is contributed by the air releases. Therefore, the air releases hardly seem capable of elevating the median of the dose. To look at it another way, the mean value of the air release dose is 0.469 millisieverts. If this mean value of air dose is added to the water doses, it will not substantially affect the median total dose. Although the fraction of the air dose increases substantially for lower total doses, the highest it ever gets is a little over 0.4. For most values, the fraction is less than 0.1 and for many it is less than 0.05. This is not the answer.

Table 12-8 Statistics on Total Effective Dose Equivalent for Different Receptors

Family	Family Member	Point Estimate (mSv)*	Run 0 (mSv)*	Mean (mSv)*	Median (mSv)*
Delivery Family	Adult Female	6.106	6.091	12.812	9.072
Delivery Family	Adult Male	6.283	6.266	13.001	9.281
Delivery Family	Child Born 1955	5.200	5.180	10.164	7.993
Delivery Family	Child Born 1964	2.090	2.081	4.498	3.159
Migrant Family	Adult Female	0.447	0.438	0.793	0.562
Migrant Family	Adult Male	0.624	0.614	1.117	0.756
Migrant Family	Child Born 1955	2.178	2.160	3.676	2.489
Migrant Family	Child Born 1964	0.083	0.081	0.127	0.093
Near Water Family	Adult Female	2.091	2.057	3.431	2.738
Near Water Family	Adult Male	2.205	2.170	3.574	2.929
Near Water Family	Child Born 1955	3.137	3.099	4.815	4.311
Near Water Family	Child Born 1964	1.759	1.734	2.850	2.290
Outdoor Family	Adult Female	3.030	3.001	4.687	4.263
Outdoor Family	Adult Male	4.216	4.169	6.055	5.546
Outdoor Family	Child Born 1955	9.435	9.383	13.331	10.988
Outdoor Family	Child Born 1964	1.826	1.810	2.951	2.309
Rural Family One	Adult Female	0.303	0.299	0.502	0.387
Rural Family One	Adult Male	0.423	0.418	0.712	0.538
Rural Family One	Child Born 1955	1.589	1.580	2.681	1.697
Rural Family One	Child Born 1964	0.072	0.071	0.093	0.077
Rural Family Two	Adult Female	0.696	0.685	1.174	0.890
Rural Family Two	Adult Male	0.974	0.961	1.655	1.198
Rural Family Two	Child Born 1955	3.751	3.729	6.362	4.006
Rural Family Two	Child Born 1964	0.140	0.137	0.190	0.153
Urban Family	Adult Female	0.330	0.325	0.447	0.284
Urban Family	Adult Male	0.731	0.723	0.895	0.698
Urban Family	Child Born 1955	2.686	2.675	4.314	2.551
Urban Family	Child Born 1964	0.107	0.106	0.115	0.107

Table 12-9 Comparison of Mean Dose and Median Dose with Point-Estimate Dose for each Receptor

Family	Family Member	Pt. Estimate (mSv)*	Mean (mSv)*	Ratio, Mean over Point Estimate	Median (mSv)*	Ratio, Median over Point Estimate
Delivery Family	Adult Female	6.106	12.812	2.098	9.072	1.486
Delivery Family	Adult Male	6.283	13.001	2.069	9.281	1.477
Delivery Family	Child Born 1955	5.200	10.164	1.955	7.993	1.537
Delivery Family	Child Born 1964	2.090	4.498	2.153	3.159	1.512
Migrant Family	Adult Female	0.447	0.793	1.773	0.562	1.255
Migrant Family	Adult Male	0.624	1.117	1.790	0.756	1.211
Migrant Family	Child Born 1955	2.178	3.676	1.688	2.489	1.143
Migrant Family	Child Born 1964	0.083	0.127	1.524	0.093	1.119
Near Water Family	Adult Female	2.091	3.431	1.641	2.738	1.309
Near Water Family	Adult Male	2.205	3.574	1.621	2.929	1.328
Near Water Family	Child Born 1955	3.137	4.815	1.535	4.311	1.374
Near Water Family	Child Born 1964	1.759	2.850	1.621	2.290	1.302
Outdoor Family	Adult Female	3.030	4.687	1.547	4.263	1.407
Outdoor Family	Adult Male	4.216	6.055	1.436	5.546	1.316
Outdoor Family	Child Born 1955	9.435	13.331	1.413	10.988	1.165
Outdoor Family	Child Born 1964	1.826	2.951	1.616	2.309	1.264
Rural Family One	Adult Female	0.303	0.502	1.656	0.387	1.277
Rural Family One	Adult Male	0.423	0.712	1.683	0.538	1.270
Rural Family One	Child Born 1955	1.589	2.681	1.687	1.697	1.067
Rural Family One	Child Born 1964	0.072	0.093	1.295	0.077	1.067
Rural Family Two	Adult Female	0.696	1.174	1.688	0.890	1.279
Rural Family Two	Adult Male	0.974	1.655	1.700	1.198	1.231
Rural Family Two	Child Born 1955	3.751	6.362	1.696	4.006	1.068
Rural Family Two	Child Born 1964	0.140	0.190	1.362	0.153	1.094
Urban Family	Adult Female	0.330	0.447	1.353	0.284	0.859
Urban Family	Adult Male	0.731	0.895	1.224	0.698	0.954
Urban Family	Child Born 1955	2.686	4.314	1.606	2.551	0.950
Urban Family	Child Born 1964	0.107	0.115	1.068	0.107	0.996

*mSv = millisieverts.

The components of the water dose seem to provide a more plausible explanation of the behavior of uncertain doses. For each realization, the doses resulting from cesium-137, strontium-90, and phosphorus-32 were listed as well as their sum. The median of the "sum" is 8.419 millisieverts. This sum accounts for most of the water-dose median (8.419/8.539 = 0.986) and most of the total-dose median (8.419/9.072 = 0.928). The median of doses for each isotope were found separately, and the sum of these medians is 5.857 millisieverts. This is much closer to the point-estimate dose of 6.106 millisieverts. This appears to suggest that if the three isotopic components of dose were summed separately, they would correspond closely to the base case. However, because these doses are summed for each realization, the sum is, on average, larger than the point estimate. Because the uptake factors for these three isotopes vary

independently and randomly, the total dose can be considered approximately to be the sum of three lognormal distributions. The smaller, but substantial, doses from strontium-90 and phosphorus-32 essentially make a "floor" on the dose from cesium-137 and boost the median (as shown) by a factor of about 50 percent over what would occur if the uptake factors were correlated.

In summary, it appears that the confidence intervals do not include the point estimates. This appears to be a result of both the assumptions regarding the input variable distributions and the mathematical form of the dose model. This also shows the utility of performing an uncertainty analysis, which may give a different perspective on the problem than a point estimate.

12.6 Secondary Results

The primary result of the overall uncertainty analysis was to establish distributions of dose estimates that incorporated uncertainty in the input variables and to establish the statistical confidence intervals about the mean values of those distributions. However, the large amount of data generated in performing this uncertainty analysis can be used to provide additional insights into how doses and the uncertainties in doses depend on the uncertain variables analyzed. This additional information can be useful in determining how to refine the modeling approaches or how to prioritize the need for additional site data.

In Chapter 11 dose to various receptors was explored by identifying important factors such as radionuclides, year of exposure, exposure pathways, and exposure routes. These results provided information on the structure of the model. These point estimate dose results and the conclusions drawn from them also depended on the choices made for the value of each input variables, i.e., one value for each input variable. In this chapter the dependence of dose on the distributions chosen to describe each uncertain input variable is explored; i.e., the attention here is on the dependence of dose on the uncertain input variables. These analyses are based upon the sampled input variables and results of the uncertainty analysis described in previous sections. By aligning the input variable values with the respective resultant dose values, several statistical techniques can be used to evaluate how a change in an input variable changes the output (dose).

The dependence of dose on the uncertain input variables can be considered to have two components: 1) how effective a change in an input variable is in producing a change in dose and 2) how widely a particular input variable changes. Big variations in dose may be produced in three ways: 1) dose may be very sensitive to a particular variable, so even modest variations in that input variable produce substantial variations in dose; 2) dose may be moderately sensitive to a particular variable, but the variable is very uncertain, so large variations in dose are produced; 3) dose may be very sensitive to a highly variable input, so extremely large variations in dose are produced. The two components describing the dependence of dose on uncertain input variable are sometimes characterized by two coefficients: 1) a sensitivity coefficient and 2) a variance or uncertainty coefficient. Broadly speaking, the sensitivity coefficient is the ratio of the fractional change in dose to the fractional change of an input variable. The variance or uncertainty coefficient is the fraction of uncertainty in dose attributable to the uncertainty in an input variable. Both of these aspects are explored in the Sections that follow.

Sections 12.6.1 and 12.6.2 discuss how some general statistical methods can be used to describe the relationships between dose and input variables. Scatter plots are used to give a graphical assessment of the input-dose relationship. The correlation coefficient (assuming a linear relationship) was used to quantify the relationship between dose and input variables. Section 12.6.3 describes how the variability in dose depends on input variables (i.e., variance or uncertainty considerations); Section 12.6.4 describes how the dose depends on input variables (i.e., sensitivity considerations). Only a few examples were analyzed and described here. An extensive analysis of this type is beyond the scope of the current study. More extensive analyses of this type may be appropriate in the future. The limited analyses described

here generally reinforce the understanding of the generation of doses by SRS releases obtained from the modeling and results described in previous Chapters.

The names and characteristics of the uncertain variables analyzed are described in Table 12-2. However, to facilitate this discussion, the variable names and their physical meaning are repeated in Table 12-10 here.

Table 12-10 Names of Uncertain Variables and Their Physical Meaning

Variable Name	Physical Meaning
Water Pathway	
CLBFF, Cs-137	Bioconcentration factor for Cs-137 in fish
CLBFF, Sr-90	Bioconcentration factor for Sr-90 in fish
CLBFF, P-32	Bioconcentration factor for P-32 in fish
Air Pathway	
LEAFRS	Soil resuspension factor; determines the amount of radioactivity deposited on the ground that is resuspended; affects urban and rural scenarios.
RADIUS	Radius of particles that can deposit on the soil and plant surfaces.
WTIM	Weathering rate constant determines the rate at which deposited radioactivity is removed by weathering processes from crop surfaces.
DPVRES	Deposition velocity from air to plant surfaces for resuspended activity.
DEPFR1, 2	Dry and wet deposition fraction; how much of the radioactivity deposited on plant surfaces is retained and absorbed by the plant.
GRWPA, 5&6	Growing period for animal forage; the longer the growing period, the more exposure the plant has to air deposition of radionuclides.
CLFMT I-131	Uptake of I-131 by beef muscle; transfer to meat from cattle food.
F	Release factor characterizes the uncertainty in air and water releases of radionuclides from the SRS.
CONSUM, 5	Animal forage consumption rate; scales forage consumption by milk and beef animals.
CLFMK I-131	Uptake of I-131 by milk; transfer to milk from cow food.
BIOMA2	Animal forage standing biomass (wet) describes how much grass is in the field per unit area.

12.6.1 Scatter Plots

Figure 12-5 shows examples of scatter plots for the Delivery Family Adult Female. Simple scatter plots of the dose versus a sampled input variable (e.g., CLBFF-Cs, the bioconcentration of cesium in fish) can give a quick visual display of the importance of a particular variable in affecting the dose. For example, the left and right plots are, respectively, the plot of dose versus the cesium-bioconcentration in fish and the plot of dose versus the suspended particle radius. The pattern in the left plot resembles that of a line, indicating a strong linear relationship between the uptake of cesium in fish and the resultant dose to the Delivery Family Adult Female. The much more random “snow” pattern in the right plot indicates that the dose to the Delivery Family Adult Female is not significantly affected by the resuspended soil-particle radius.

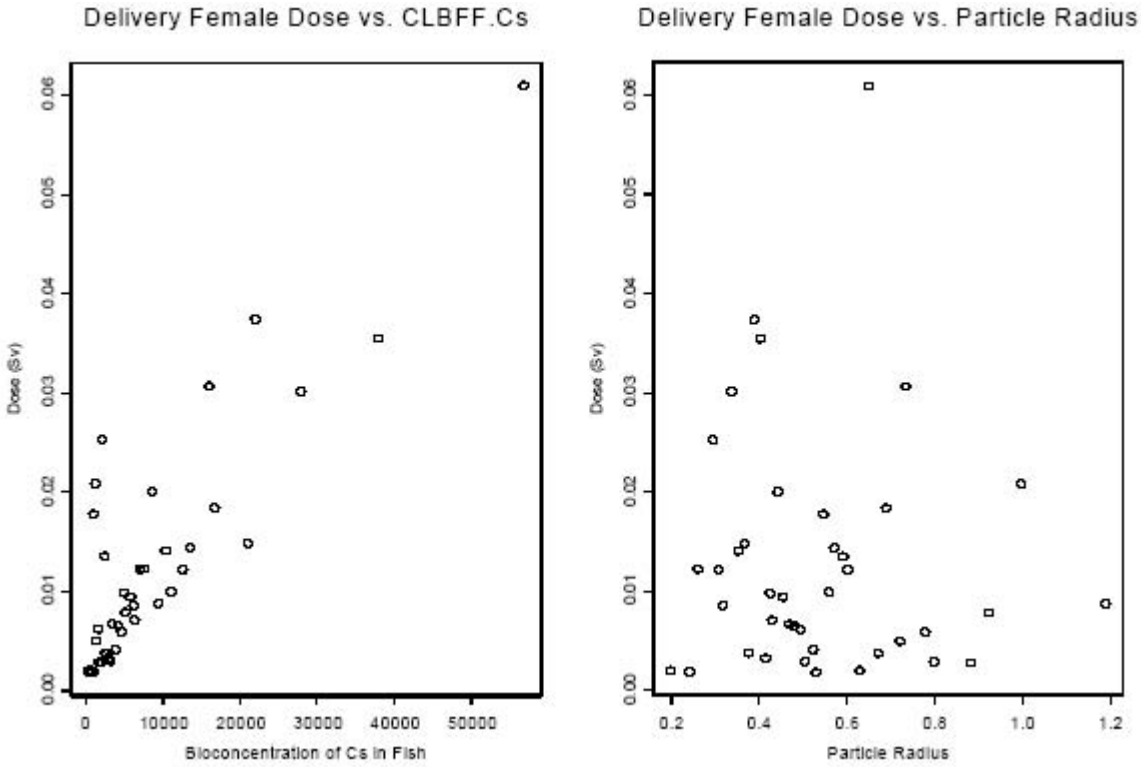


Figure 12-5 Examples of Scatter Plots for the Delivery Family Adult Female: Dose vs. Bioconcentration of Cesium in Fish (Left Plot) and Dose vs. Particle Radius (Right Plot)

12.6.2 Correlation Coefficients

Correlation coefficients are another way of expressing the relationship between input variables and the dose. Again, this analysis focused on linear relationships and did not consider nonlinear relationships such as dose as a function of the square of the particle radius. The correlation coefficient is a quantitative, statistical measure that represents the strength of the linear relation between two random variables (6). In this case, the two random variables are the dose and the input variable selected for examination. Correlation coefficients range between -1.0 and +1.0. A value of +1.0 indicates a perfect direct relationship; a value of -1.0 indicates a perfect negative relationship. For example, the size of one's bank account may have a correlation of +1 with the size of one's paycheck, but a correlation of -1 with the size of one's expenditures.

Correlation coefficients for the CLBFF-Cs and radius scatter plots given above are 0.864 and -0.032, respectively. This quantitatively demonstrates that CLBFF-Cs is over 10 times more effective in changing the dose than is the radius.

12.6.3 Stepwise Regression Analysis

The dose to the 28 hypothetical individuals considered in this analysis is the summation of incremental doses through various exposure pathways and from various radionuclides. These doses depend on many environmental and behavioral variables. Thus, in determining the most important variables, it is important to consider them together for the purpose of establishing an order of most to least significant within the

group. This can be accomplished with the stepwise regression analysis technique. In this type of analysis, the relationship of the dose to all of the sampled input variables is represented as a polynomial; that is, the statistical model for the dose is a sum of linear terms where each term is a product of a coefficient and an input variable.

A main focus of the stepwise regression analysis is to determine the best value of the coefficient for each input variable. This statistical model is built in a stepwise fashion by sequentially adding (or subtracting) one independent variable at a time. The result of this analysis is a statistical model that describes the dose as a function of a small set of independent variables (e.g., Dose ~ f(CLBFF-Cs +CLBFF-Sr)) which are the most important in affecting dose; input variables with coefficients that are relatively small are generally deleted from the model. Because this kind of analysis can be very computationally intensive, statistical software (S-Plus; Insightful, 2002) was used. As the software performed the stepwise regression, it automatically checked the coefficient of determination (R^2) and the statistical level of significance. Variables that do not have significant correlation with the dose or result in an increase in the R^2 are generally left out.

The statistic R^2, the coefficient of determination, is a "measure of the proportion of the total variability in the dependent variable that is accounted for by the regression equation" (7). A value of 1 for R^2 suggests that all the variability in the dependent variable (dose) is accounted for by the variability in the selected independent variables. A low value for R^2 is often interpreted to mean that the variability in the dependent variable is not linked to the variation in the independent variables. This is true to some extent. A low R^2 is many times a reflection of a small range of variation of the independent variable. Or, it might indicate that some of the variables should be considered in a nonlinear fashion (e.g., X^2 or 1/x).

The stepwise regression analysis was conducted for the hypothetical families that consume fish (Delivery, Near Water, and Outdoor Family) by considering all 14 of the sampled input variables. For those hypothetical families that did not consume fish (Migrant, Rural Families 1 and 2, Urban/Suburban), the stepwise regression analysis was conducted without any of the bioconcentration factors in fish (i.e., CLBFF for cesium, strontium, and phosphorus).

Table 12-10 and Table 12-11 present the results of the stepwise regression analysis, respectively, for the Delivery Family (which ate fish) and Rural Family 2 (which did not). Table 12-12 presents the stepwise regression analysis results for each of the Adult Females representing all scenarios. The five most important variables are listed in order of importance. The rank ordering is based upon the correlation coefficient and the contribution to the R^2. In general, for those variables that contribute more than 0.1 to the multiple R^2, there is little uncertainty concerning their rank. In this analysis, this generally applies to the top three variables. Variables ranked fourth and fifth made a very small contribution to the multiple R^2 and thus might change with a different sampling using the Latin Hypercube Sampling program.

Table 12-10 Delivery Family Stepwise Regression Results Listing Input Variables in Order of Influence and the Sum of R^2 for the Variables Listed

Adult Female	Adult Male	Child Born 1955	Child Born 1964
CLBFF-Cs	CLBFF-Cs	CLBFF-Sr	CLBFF-Sr
CLBFF-Sr	CLBFF-Sr	CLBFF-Cs	CLBFF-Cs
F	F	CLFMK I-131	CLBFF-P
CLBFF-P	BIOMA2	CLFMT I-131	BIOMA2
BIOMA2	CLBFF-P	LEAFRS	F
$R^2 = 0.98$	$R^2 = 0.98$	$R^2 = 0.96$	$R^2 = 0.97$

Note: Refer back to Table 12-2 for descriptions of the variables.

Table 12-11 Rural Family 2 Regression Results Listing Input Variables in Order of Influence and the Sum of R^2 for the Variables Listed

Adult Female	Adult Male	Child Born 1955	Child Born 1964
DEPRF1,2	LEAFRS	LEAFRS	LEAFRS
LEAFRS	DEPRF1,2	F	F
F	F	DEPFR 1,2	DEPFR 1,2
DPVRES	DPVRES	DPVRES	DPVRES
CONSUM	CONSUM	CONSUM	CONSUM
$R^2 = 0.597$	$R^2 = 0.576$	$R^2 = 0.793$	$R^2 = 0.5793$

Note: Refer back to Table 12-2 for full descriptions of the variables.

Table 12-12 Stepwise Regression Results for the Adult Female of Each Scenario

Delivery Female[1]	Migrant Female	Near Water Female[1]	Outdoor Female[1]	Rural Family 1 Female	Rural Family 2 Female	Urban Female
CLBFF-Cs	LEAFRS	LEAFRS	CLBFF-Cs	DEPRF1,2	DEPRF1,2	DEPRF 1,2
CLBFF-Sr	DEPFR 1,2	DEPRF 1,2	CLBFF-Sr	LEAFRS	LEAFRS	WTIM
F	F	F	CLBFF-P	F	F	CLFMK I-131
BIOMA2	DPVRES	DPVRES	DEPRF1,2	DPVRES	DPVRES	CLMFT I-131
CLBFF-P	CONSUM	CONSUM	CLFMT I-131	CONSUM	CONSUM	CONSUM
$R^2 = 0.98$	$R^2 = 0.692$	$R^2 = 0.692$	$R^2 = 0.939$	$R^2 = 0.593$	$R^2 = 0.597$	$R^2 = 0.731$

[1] Ate Fish
Note: Refer back to Table 12-2 for full descriptions of the variables.

In the case of the Delivery Family (Table 12-11), the most important variables to the dose are the bioconcentration factors for cesium, strontium, and phosphorus in fish, suggesting that much of the variability in dose to the Delivery Family is from the variability of the uptake factors for cesium, strontium, and phosphorus by fish. As described in Chapter 3, the Delivery Family gets 50 percent of its fish from Lower Three Runs Creek, which drains from the area with P and R reactors. This list of variables may be shortened even more by dropping CLBFF-P, because CLBFF-Cs and CLBFF-Sr yield an R^2 of 0.91, which is most of the variation.

After strontium and cesium in fish, the important variables for the Delivery Family Child Born in 1955 are those variables associated with uptake of iodine-131by beef cattle and milk cows (CLFMT I-131 and CLFMK I-131), reflecting the relatively large releases of iodine-131 in the early years of site operation.

The most important variables for the Rural Family 2 (Table 12-12) are related, as expected, to ingestion. However, the variable LEAFRS (resuspension factor) could also impact inhalation dose. The variable CONSUM (forage consumption rate) is associated with the uptake of radionuclides (especially I-131) by beef cattle and milk cows. The R^2 values for this set of stepwise regression analyses (approximately 0.6

for each receptor) indicate that the relationship between the independent variables and the dose may not be linear or that there is a small range for one of the variables (CONSUM has a range of 9 to 62; less than 1 order of magnitude).

Table 12-13 compares the stepwise regression results for the Adult Female in each of the seven scenarios; this is intended to show how results are scenario dependent, but comparisons based on other family members may identify different variables. In general, the results of the stepwise regression analysis are consistent with the scenarios as they are defined. Specifically, of the hypothetical families that eat fish, the Delivery Family got its fish from Lower Three Runs Creek and the Savannah River, which exposed them to cesium, strontium, and phosphorus releases from the SRS at relatively higher concentrations than those families obtaining fish from only the Savannah River. The members of the Outdoor Family fished along the Savannah River shoreline adjacent to the SRS and were thus exposed to radionuclides transported through the wetlands area. Thus, it is reasonable that the most significant variables in determining the variation in dose to these families are the variations in bioconcentration factors for cesium and strontium.

In contrast, although the Near Water Family got its fish from the Savannah River, the consumption of fish was smaller than for the Outdoor Family. The bioconcentration factors cesium, strontium, and phosphorus in fish were not among the most important variables for this scenario. Instead, the Near Water and Migrant families exhibit a similar ranking of important variables. The most important variables for these families are related to exposure by air and consumption of locally grown fruits, vegetables, milk and beef. Although these families lived on opposite sides of the SRS, the similarity in the sensitivity rankings suggest that exposure through pathways related to air transport did not vary substantially around the circumference of the SRS.

The two Rural families showed the same sensitivities to input variables. Like the Migrant and Near Water families, the most important variables affecting variation in dose are those related to airborne transport and uptake through farm products.

The first-ranked variable for the Migrant and Near Water families is LEAFRS. This variable contributed 0.41 to the total R^2. The second-ranked variable DEPRF1,2 contributed 0.15. By contrast, the contribution to the total R^2 for the two Rural families was 0.23 for DEPFR1,2 and 0.21 for LEAFRS. The reasons for these differences in variable ranking and their relative contribution are not clear.

The ranking of variables for the Urban/Suburban Family is different from those families living in more rural settings closer to the SRS. This family's primary exposure to radionuclides from the site was through the milk produced in the New Ellenton area. DEPRF1,2 is the most significant variable with a contribution of 0.4 to the total R^2. Other variables are for removal of deposited radionuclides from plant surfaces (WTIM) and the uptake of I-131 into milk and beef (CLFMT I-131 and CLFMK I-131).

Several scenarios show sensitivity to the release factor, an expected result: the more released, the higher the dose.

12.6.4 Sensitivity Coefficients

In Section 12.6.3, the most important variables were ranked in order of their contribution to the coefficient of determination, R^2. This ranking actually is for the contribution to the variability in the dose, not the sensitivity, per se. To rank the variables in terms of their sensitivity coefficient (how a fractional change in the dose reflects a fractional change in the input variable), the uncertain variables were transformed to dimensionless quantities by dividing by the mean of each respective distribution. The transformed variables were used in the stepwise regression analysis, and while the variables found to be

dominant remained the same, the ranking changed. In this analysis, the sensitivity coefficients are based on the slope of a least-squares-fit between dose and the input variable.

Table 12-13 presents the variable rankings and their sensitivity coefficients for the Delivery Family and Rural Family 2 Adult Females. These rankings are related to how important the variables are in determining the dose; these rankings reflect only slightly the degree of uncertainty in the input variables themselves.

Table 12-13 Variable Rankings and Their Sensitivity Coefficients

Delivery Family Adult Female		Rural Family 2 Adult Female	
Variable	**Sensitivity Coefficient**	**Variable**	**Sensitivity Coefficient**
F	0.66	F	0.92
BIOMA2	-0.65	DEPFR 1,2	0.84
CLBFF-Cs	0.64	CONSUM	0.38
CLBFF-Sr	0.13	LEAFRS	0.15
CLBFF-P	0.01	DPRVES	0.09

Note: Change in sign for BIOMA2 that is more consistent with the analytical model.

12.7 Summary

The following are some of the key aspects of the uncertainty analysis:

- Forty realizations were used to investigate the uncertainty behavior of dose for each of 28 hypothetical receptors.
- In general, the distribution of dose for each receptor behaved similar to a lognormal distribution.
- Consistent with the shape of a lognormal distribution, the mean of each dose distribution was higher than the median of the distribution.
- The confidence intervals were estimated for the mean and median. The sizes of these intervals were not large compared to the corresponding mean and median. This appears to reflect the sample size (40) and the underlying uncertainty quantified for the dose distributions. Also, modeling uncertainty was not estimated; only uncertainty related to the input variables was quantified.
- In general, the mean and median for the uncertainty analysis are larger than the corresponding point estimate of dose. This appears to reflect the interaction among the uncertain variablesand the complexity of the dose model.
- The calculated confidence intervals do not contain the point estimates of dose for the reasons cited in the previous bullet.
- The sensitivity analysis generally shows the same dominance of the milk and beef ingestion pathways for those scenarios dominated by dose from air release (as did the point-estimate analysis). Similarly, for doses from water releases, the fish ingestion dominates—specifically, the fish bioconcentration factors for cesium-137, strontium-90, and phosphorus-32.

12.8 References

1. Iman, R.L. & M.J. Shortencarier (1984, March). A FORTRAN 77 program and user's guide for the generation of latin hypercube and random samples for use with computer models. NUREG/CR-3624 (SAND83-2365). U.S. Nuclear Regulatory Commission. Washington, DC.

2. Decision Engineering, Inc. (2000). Crystal Ball 2000, user manual.

3. Palisade Corporation (2002, February). @RISK, advanced risk analysis for spreadsheets. Version 4.5.

4. Till, J.E. & H.R. Meyer, (eds.) (1983, September). Radiological Assessment, A Textbook on Environmental Dose Analysis. NUREG/CR-3332. ORNL-5968.

5. Bowen, W.M. & C.A. Bennett (1988, December). Statistical methods for nuclear material management. NUREG/CR-4604. U.S. Nuclear Regulatory Commission. Washington, DC.

6. Lurie, D. & R.H. Moore (1994, December). Moore RH. Applying statistics. NUREG-1475. U.S. Nuclear Regulatory Commission. Washington, DC.

7. Mac, Berthouex P., & L.C. Brown (1994). Statistics for environmental engineers. Lewis Publishers. Boca Raton, Florida.