Recommended Practice Regarding Selection, Application, and Interpretation of Sensitivity Analysis Methods Applied to Food Safety Process Risk Models

Prepared by:

H. Christopher Frey
Amirhossein Mokhtari
Junyu Zheng
Computational Laboratory for Energy, Air, and Risk
Department of Civil, Construction, and Environmental Engineering
North Carolina State University
Raleigh, NC 27695-7908

Prepared for:

Office of Risk Assessment and Cost-Benefit Analysis
U.S. Department of Agriculture
Washington, DC

January 30, 2004

Acknowledgments

This report was prepared as part of Cooperative Agreement 58-0111-0-005 between North Carolina State University and the Office of Risk Assessment and Cost Benefit Analysis (ORACBA) of the U.S. Department of Agriculture (USDA).

We are grateful to the following individuals who provided input and comments regarding this guidance document:

John Bowers, U.S. Food and Drug Administration
Clark Carrington, U.S. Food and Drug Administration
Peter Cowen, North Carolina State University
Eric Ebel, U.S. Department of Agriculture
Aamir Fazil, Health Canada
Lee-Ann Jaykus, North Carolina State University
Greg Paoli, Decisionalysis, Inc.
Mark Powell, U.S. Department of Agriculture
Wayne Schlosser, U.S. Department of Agriculture
Mark Walderhaug, U.S. Food and Drug Administration

Disclaimers

Any opinions, findings, conclusions, or recommendations expressed in this guidance document are those of the authors and do not necessarily reflect the views of the U.S. Department of Agriculture.

Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

The purpose of this document is to provide guidance to food safety risk analysis practitioners regarding the application of sensitivity analysis to food safety process risk models, particularly those for food-borne microbial pathogens. The purpose of such models is typically to provide insight regarding possible risk management strategies. Sensitivity analysis is a methodology for identifying the key model inputs that contribute the most to variation in a selected model output or to the highest values of the output. Thus, an analyst is often interested in identifying which one or subset of potentially many controllable variables is of greatest importance with regard to human exposure and risk. Other reasons for performing sensitivity analysis include assisting in the process of model development and prioritizing additional work aimed at reducing uncertainty in an assessment.

The key questions that are addressed by this document are listed in Section 1.1. Section 1.2 provides background on recent work that has lead to the creation of this document. Section 1.3 provides an overview of the document aimed at assisting the reader in quickly accessing information regarding a particular topic.

1.1 Key Questions

This document assists the practitioner with regard to the following key questions:

- When should I perform sensitivity analysis? (Chapter 2)
- How do I prepare a model to facilitate sensitivity analysis? (Chapter 3)
- What are key considerations in the development of scenarios that are the basis for sensitivity analysis? (Chapter 4)
- What are some typical sensitivity analysis methods, and how can I select among them? (Chapter 5)
- How should particular sensitivity analysis methods be applied? (Chapter 6)
- How should the results of sensitivity analysis be presented and interpreted? (Chapter 7)

The organization of the document with respect to each of these key questions is illustrated in Figure 1-1.

This document includes a glossary of common terms that are used with respect to sensitivity analysis and food safety process risk assessment modeling.

1.2 Background

There has been increasing development and use of quantitative models for food safety process risk assessment. These models can be large and complex. Thus, it can be difficult to prioritize controllable variables that are most promising with respect to risk management goals. As a matter of good practice, it is important to evaluate how a model responds to changes in its inputs as part of the process of model development, verification, and validation. Moreover, insight regarding key sources of uncertainty in a model can be used to prioritize additional data collection or research in order to reduce uncertainty.

There are several techniques for sensitivity analysis used by practitioners and analysts in numerous fields, including microbial risk assessment. The most commonly used methods are those that are built-in features of a particular software tool. An example is the use of sample or rank correlation coefficients in software packages such as Crystal Ball™ or @Risk™. However, there are other sensitivity analysis methods, including some used in other fields, that may be useful if applied to food safety process risk models.

This document is the product of Phase 3 of a three year, three phase project. An objective of this project is to transfer, apply, and adapt sensitivity analysis methods developed in other disciplines (e.g. complex engineering systems) to food-safety risk assessment. This work has been conducted under a cooperative agreement between the U.S. Department of Agriculture (USDA) and North Carolina State University (NCSU). Phase 1 involved the identification and preliminary evaluation of promising sensitivity analysis methods, including literature review and input from experts and practitioners. Phase 2 involved an intensive evaluation of approximately a dozen sensitivity analysis methods applied to two food safety process risk models.

In Phase 1, NCSU prepared a literature review of sensitivity analysis methods and convened a workshop of sensitivity analysis experts and food safety process risk model practitioners. The workshop was held June 11-12, 2001 at NCSU. The workshop addressed three key issues pertaining to the application of sensitivity analysis in food safety risk assessment including: (1) key criteria for sensitivity analysis methods applied to food safety models; (2) identification of the most promising sensitivity analysis methods for application to food safety process risk models; and (3) key needs for implementation and demonstration of sensitivity analysis methods. The workshop participants agreed that different methods of sensitivity analysis

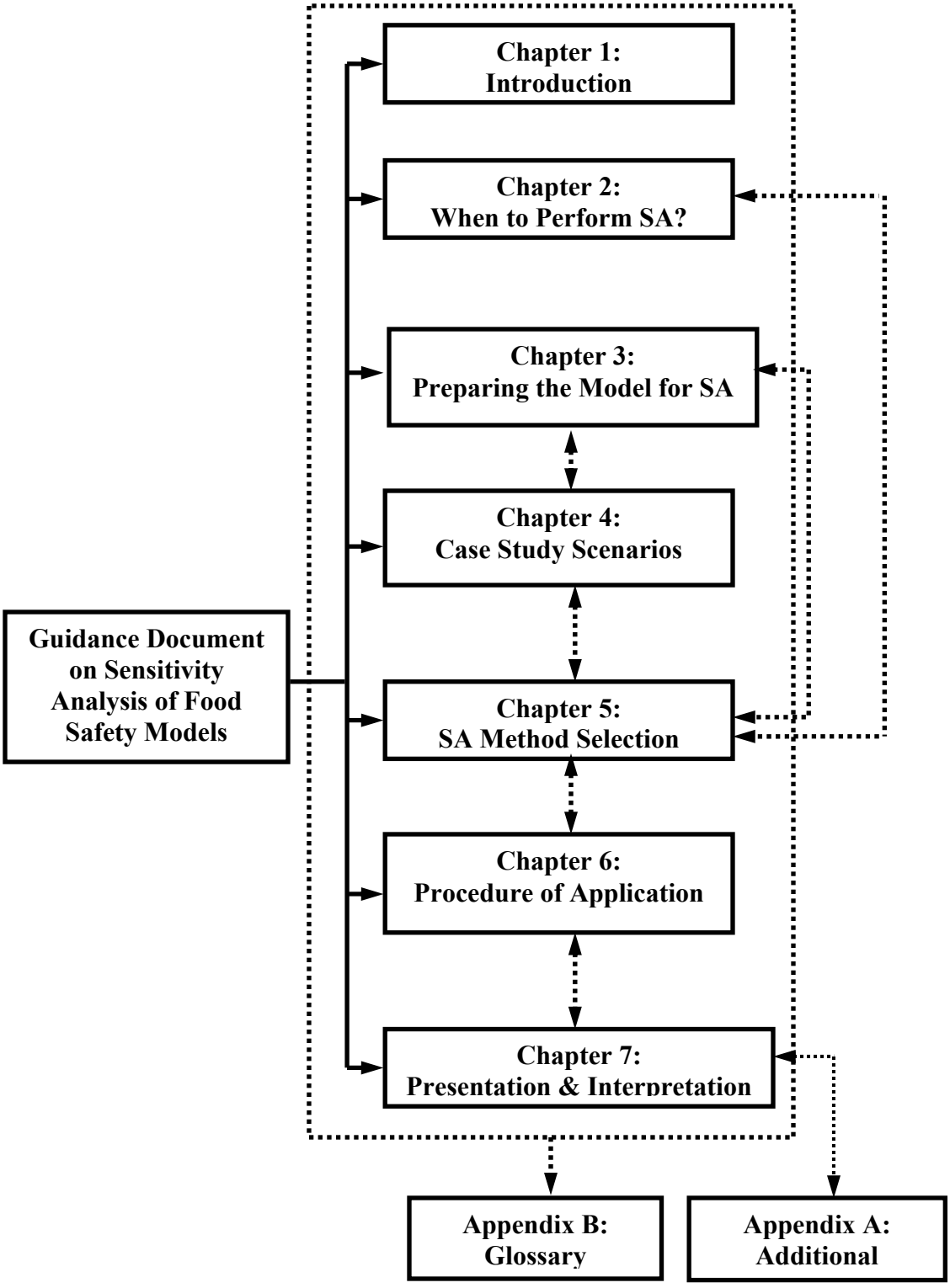

Figure 1-1. Flow Diagram of the Organization of the Guidance Document.

should be evaluated based upon application to more than one food safety risk model. The process of testing methods would help establish a track record for specific methods applied to food safety process risk models. The workshop also recommended that a guidance document be developed to assist practitioners with regard to the selection of sensitivity analysis methods, their application, interpretation, and reporting.

As a response to the workshop recommendations, in Phase 2 NCSU evaluated several sensitivity analysis methods based upon application to two food safety risk assessment models: (1) *Listeria monocytogenes* in Ready-to-Eat (RTE) foods; and (2) *E. coli O157:H7* in ground beef (Frey *et. al.,* 2003). One of the key outputs of the Phase 2 report is a summary table comparing the capabilities of many sensitivity analysis methods with regard to the characteristics of food safety process risk models and with regard to analytic objectives. Therefore, the Phase 2 report provides an experiential basis for the recommendations contained in this document.

In March 2003, the draft Phase 2 report and a preliminary outline of this document were reviewed by an international workshop comprised of food safety risk practitioners. By consensus, the key objectives for this guidance document were identified as follows:

- Develop guidelines that are not too prescriptive, but that provide useful boundaries and principles for selecting, using, and interpreting results from sensitivity analysis methods
- Account for scenario uncertainty, model uncertainty, and model input uncertainty and variability
- Define terminology
- Identify and compare selected sensitivity analysis methods
- Identify modeling requirements to facilitate sensitivity analysis, and recommend approaches for design and implementation of models.

In the next section, the role of this document, as envisioned by the participants of the Phase 1 and 2 workshops, is briefly described.

1.3 The Role of This Document

Based upon the recommendations of the June 2001 workshop during Phase 1, and confirmed by the March 2003 workshop during Phase 2, there is a need for guidance to practitioners regarding selection, application, interpretation, reporting, and documentation of sensitivity analysis methods. Considering the wide range of sensitivity analysis methods, a

practitioner should clearly understand which methods are appropriate for a specific application. A need was expressed regarding guidance on procedures for application for each method in order to substantially facilitate sensitivity analysis of a model. A key step of the analysis is to interpret and present the results of the analysis based on each method. Thus, this guidance document is intended to help practitioners in selecting specific sensitivity analysis methods that are relevant to a particular application and to the characteristics of the model. Practitioners can use this document to aid in interpreting results from a sensitivity analysis in response to a particular modeling objective. Examples of modeling objectives include: prioritizing controllable sources of variability in exposure and risk in order to develop risk management recommendations; identification of key sources of variability and uncertainty in order to facilitate model development, verification, and validation; or prioritization of key sources of variability and uncertainty in order to prioritize additional data collection and research.

The following paragraphs briefly summarize the content of the remaining chapters.

When should I perform sensitivity analysis? Chapter 2 discusses several key motivations for performing sensitivity analysis, including prioritization of critical control points, identification and prioritization of key sources of uncertainty and variability in order to prioritize data collection and research, and model refinement.

How do I prepare a model to facilitate sensitivity analysis? Chapter 3 discusses the key considerations in developing or preparing a model in order to facilitate sensitivity analysis. Key issues include model structure, identification and accessing of model inputs, decision-based selection of model outputs, specification of the type of probabilistic simulation to be performed, and suggestions for how to modify an existing model. Recommendations also are made regarding the development of new food safety process risk models. These recommendations involve modeling environment, determination of the type of probabilistic simulation, modeling strategies, and documentation of models.

What are key considerations in the development of scenarios that are the basis for sensitivity analysis? Chapter 4 discusses key concepts in defining a case scenario as the basis of sensitivity analysis. Key concepts include identification of susceptible subpopulations, identification of exposure pathways of interest and selected food categories, spatial and temporal dimension of the analysis, and the probabilistic features of sensitivity analysis. Clear specification of these issues in a scenario makes the analysis transparent.

What are some typical sensitivity analysis methods and how can I select among them? Chapter 5 overviews typical sensitivity analysis methods and provides guidance in selection of an appropriate method for a case scenario, considering factors such as model characteristics, objectives or purposes of an analysis, computing resources available and simulation properties. Some decision trees also are presented to help practitioners in the process of selecting appropriate sensitivity analysis methods.

How should particular sensitivity analysis methods be applied? Chapter 6 introduces procedures for application of selected sensitivity analysis methods to food safety process risk models. Procedures presented in this chapter do not depend on application of specific statistical software package.

How should the results of sensitivity analysis be presented and interpreted? Chapter 7 discusses how to present and interpret sensitivity analysis results for selected methods in Chapter 5. For each method, an example is provided illustrating the presentation and interpretation of the results of sensitivity analysis. The issue of audience also is discussed briefly in this chapter. For different audiences, the ways for presenting and interpreting the sensitivity analysis results may be different.

Appendix *A* contains summary information on selected less commonly applied but potentially useful sensitivity analysis methods in addition to those addressed in the main text.

Appendix *B* contains a glossary of technical terms to assist the reader.

2 GUIDANCE ON WHEN TO PERFORM SENSITIVITY ANALYSIS

The objective of this chapter is to provide guidance to identify situations in which sensitivity analysis of a food safety process risk model is useful. Saltelli (2000) defines sensitivity analysis as the study of how the variation in the output of a model can be apportioned, qualitatively or quantitatively, among model inputs. Similarly, Cullen and Frey (1999) define sensitivity analysis as the assessment of the impact of changes in input values on model outputs. Sensitivity analysis of risk assessment models can be used to identify the most significant inputs governing exposure or risk as an aid in developing priorities for risk mitigation and management strategies. Sensitivity analysis can be used as an aid in identifying the importance of uncertainties in the model for the purpose of prioritizing additional data collection or research. Similarly, sensitivity analysis can be used to assess key sources of variability and uncertainty. Sensitivity analysis is useful for providing insight regarding model verification and regarding the robustness of model results when making decisions (Cullen and Frey, 1999).

The overall assessment objectives should be clearly laid out before the model is built. Once the objectives are defined, the model must be designed to address these objectives. For example, if the objective is to develop insight into possible risk management strategies, variables that can be controlled and variables related to risk management objectives must be incorporated into the model. Next, the answers sought from application of sensitivity analysis should be clearly listed. The usefulness of sensitivity analysis can then be assessed based on whether the available methods of sensitivity analysis can address the questions under consideration in a manner that is appropriate to the characteristics of the model.

This chapter summarizes the situations in which sensitivity analysis of the model is recommended. These situations include: (1) prioritization of potential critical control points; (2) identification of key sources of uncertainty and variability; (3) model refinement, verification, and validation; and (4) conditional analysis of the model. Conditional analysis includes “what-if” scenario analysis and identification of factors contributing to high exposure or risk. Each of these situations is briefly discussed in Sections 2.1 through 2.4, respectively. The terms and concepts specific to each of these situations are discussed in the following sections. Section 2.5 presents a summary of the chapter.

2.1 Prioritization of Potential Critical Control Points

A critical control point (CCP) is defined as a point, step, or procedure at which control can be applied, and a food safety hazard can be prevented, eliminated, or reduced to an acceptable level (Hulebak and Schlosser, 2002; Seward, 2000). Examples of CCPs include cooking, chilling, prevention of cross contamination, and product formulation controls.

The identification of CCPs is a pre model-building step in risk assessment. If the model does not include variables that represent possible CCPs, then it will not be possible to evaluate the omitted CCPs when performing sensitivity analysis. Sensitivity analysis helps prioritize the possible CCPs that are incorporated into the model. Furthermore, sensitivity analysis can provide insight regarding critical limits for a particular CCP in order to develop preventive measures. A critical limit is defined as a criterion that must be met for each preventive measure associated with a CCP. An example of a critical limit is the maximum allowed food storage temperature.

CCPs are most often based on process parameters, such as temperature, time, physical dimensions, humidity, moisture level, water activity, pH, acidity, and salt concentration. In the case that the CCP is a measurable variable in food production, handling, or preparation, the insight regarding the critical limit is of direct risk management relevance.

2.2 Identification of Key Sources of Uncertainty and Variability

Uncertainty is lack of knowledge regarding the true value of a quantity. Variability refers to real differences in values of a quantity among members of a population. Variability is also interpreted to refer to the certainty that different individuals have different exposures and different risks. In risk assessment, it is often the case that the population distribution of variability in exposure or risk is not known. Based upon available data and models, inferences are made regarding the unknown population distribution. There is uncertainty regarding the true distribution of inter-individual variability in exposure and risk. Both variability and uncertainty can be addressed quantitatively using probability distributions. In some models, a distinction is made between variability and uncertainty using a "two-dimensional" probabilistic simulation framework. In such frameworks, there is uncertainty regarding any statistic of the distribution for inter-individual variability. For example, there is uncertainty regarding the mean exposure, and regarding the exposures for any selected percentile (e.g., 50^{th} percentile, 95^{th} percentile) of the exposed population.

In order to prioritize data collection activities, it is useful to prioritize the key sources of uncertainty and variability. In many cases, the variability in exposures is influenced by only a subset of the model inputs that are subject to variability. Similarly, the uncertainty in a selected model output may be influenced by only a subset of the model inputs that are subject to uncertainty. It would be a poor allocation of scarce resources to spend an equal amount of time developing probability distributions for all model inputs, if the output is sensitive to only a small number of inputs. Sensitivity analysis can be applied to a probabilistic risk assessment model to provide insight regarding which model inputs contribute the most to uncertainty, variability, or both, for a particular model output. This insight can then be used to allocate scarce resources preferentially to data collection or research for those inputs that matter the most to the assessment.

In the case of uncertainty, the collection of additional data collection or research is the only viable method for reducing uncertainty. Because uncertainty results from lack of knowledge, it is necessary to increase the state of knowledge in order to reduce uncertainty.

In the case of variability, the collection of additional data can be used to develop more accurate estimates of variability and potentially to identify subpopulations that could be stratified into separate components of an analysis. Acquisition of data with better quality (e.g., improved representativeness) based on key sources of variability can reduce uncertainty about potential bias in the most important variable inputs. If better quality data are not available, data acquisition might only increase the precision of the potentially biased variability estimates if the data are not representative.

In some cases, it may not be feasible to collect additional data. In these situations, sensitivity analysis can provide insight regarding the robustness of the model output with regard to variation in a model input, whether due to uncertainty, variability, or both. For example, uncertainty in model inputs typically leads to uncertainty in model outputs. For a given specification of critical limits for CCPs, there may be uncertainty regarding whether a particular risk management objective can be achieved. The selection of critical limits for possible CCPs can be based upon a probabilistic criterion for the risk management objective. For example, the critical limits can be selected so that the probability of exceeding a particular exposure level for a given percentile of the population is less than five percent. Thus, sensitivity analysis can be used

to evaluate how robust the risk estimates and management strategies are to model input assumptions (Frey and Patil, 2002).

2.3 Model Refinement, Verification, and Validation

During the process of developing or refining a model, sensitivity analysis can be used as a confidence building measure with regard to model credibility. Quantitative sensitivity analysis is increasingly invoked for verification and validation of model-based analysis (Saltelli, 2002a).

Sensitivity analysis can be helpful in verification and validation of a model. Both verification and validation are important parts of quality assurance of a model. Verification is a process of checking whether a model is implemented as it was intended to be. In contrast, validation typically includes comparison of model predictions to independent empirical data for the model output under a known set of conditions.

Sensitivity analysis is useful for model verification. In particular, the objective of sensitivity analysis applied to model verification is to assess whether the model output responds appropriately to a change in model inputs. For example, if storage temperature increases, one typically expects an increase in the growth of microbial pathogens and, therefore, in exposure. If a model responds in an unacceptable way to changes in one or more inputs, then trouble-shooting efforts can be focused to identify and correct the source of the problem.

Sensitivity analysis can assist in the validation process. When validating a model, it is typically necessary to specify input values to reproduce a particular scenario for which real-world data are available for the model output. However, in many cases, the corresponding real world values are not known for all of the model inputs. Thus, judgments often must be made regarding what values to assign to some model inputs for which empirical data are not readily available. In order to determine how accurately a particular input must be estimated, it is useful to understand how substantially the model output responds to it. For example, if a model output responds by only one percent to a 50 percent change in a particular input, then it may not be important to have an accurate estimate for that particular input. In contrast, if a model output varies by 50 percent if a particular input changes by only one percent, then it could be critically important to specify an accurate value for that input as part of a validation exercise. A glaring contradiction in the relative importance of inputs from the analysis and the real world understanding indicates possible faults in the modeling.

Sensitivity analysis is helpful not only as a critique to model development as part of verification and validation, but also to guide model development. For example, when a model is developed by a team or in response to suggestion from multiple stakeholders, it is often the case that many features incorporated into the model are not essential. The identification of inputs that are of insignificant importance to the variation in the output could be used to guide the elimination of particular inputs or components of the model. Critical evaluation and reduction of the size of the model can help in preventing the model from becoming so large and unwieldy that it is no longer practical. For example, there is a tendency when building a model to include many features in response to comments, but if too many useless features are incorporated the computational resources increase with little meaningful benefit.

2.4 Conditional Analysis of the Model

Sensitivity analysis can be used for conditional analysis of a model. Conditional analysis features "what-if" scenario analysis of a model and can focus on identification of factors contributing to high exposures and risks. In "what-if" scenario analysis, specific goals with respect to the risk mitigation can be modeled. Sensitivity analysis provides a tool to evaluate how these goals can be achieved by identifying key inputs and model assumptions contributing most to the predefined scenario. Through this approach, the analysis can be framed in a way that is more responsive to the public's concerns and interests, thereby facilitating public review of the analysis. Furthermore, sensitivity analysis can provide explicit insight into the combination of key values and/or ranges of inputs that lead to the worst, or best, outcomes. The identification of worst case scenarios with respect to the exposure or risk is important for identifying possible approaches to mitigate exposure and/or risk.

As an example of a "what-if" scenario, risk managers may be interested in decreasing the mean exposure level by a specific amount. The "what-if" scenario may be in the form "what if we want to decrease the mean exposure level by one log?" Sensitivity analysis allows risk assessors and managers to consider cases where the risk can be reduced by controlling the input domains. For example, in order to constrain the mean growth of the pathogen to an acceptable level with respect to a particular management strategy it may be necessary to hold storage temperature below a specific limit, while holding storage time within a specific range. Deviation of storage temperature and storage time from the identified domains may lead to a substantial

increase in the number of pathogen organisms due to growth. Thus, constraining the inputs to specific ranges can keep the risk within acceptable levels.

2.5 Summary

This chapter briefly discussed situations in which the application of sensitivity analysis to the food safety process risk models is recommended. Sensitivity analysis is useful in: prioritizing CCPs; specifying critical limits for CCPs; identifying key sources of variability and uncertainty; refinement, verification, and validation of a model; and conducting "what-if" scenario analysis of a model. The next chapter provides recommendations regarding how to prepare food safety process risk models in order to facilitate the application of sensitivity analysis.

3 GUIDELINES FOR PREPARATION OF EXISTING OR NEW MODELS TO FACILITATE SENSITIVITY ANALYSIS

Prior to application of any sensitivity analysis methods, the model under study should be prepared for the analysis. In particular, it is important that the model is coded in a manner such that the inputs and outputs are clearly identifiable and accessible. Furthermore, the characteristics of the model, such as modularity, binning, aggregation, and probabilistic simulation options, may constrain the use of particular sensitivity analysis methods. Sensitivity analysis should be included in the list of primary modeling objectives at the time of model development. The implementation of specific model development strategies will facilitate sensitivity analysis. For an existing model, the practitioner is typically interested in applying sensitivity analysis with minimum modification to the model. However, in some situations, if the model has not been designed to facilitate sensitivity analysis, substantial modifications may be required (e.g., Patil and Frey, 2003).

Generally, a thorough understanding of the model and its limitations is essential to select well-suited sensitivity analysis methods and to determine the scope of sensitivity analysis application. The scope of sensitivity analysis may include the entire model or could be focused on specific modules or parts of a model. For example, an analyst may focus sensitivity analysis on the exposure module, which is a typical part of food safety process risk models.

This chapter focuses on the preparation of food safety process risk models for sensitivity analysis. These models are classified into two categories; (1) existing models; and (2) new models. For existing models, discussed in Section 3.1, the key steps prior to application of sensitivity analysis methods are given in Sections 3.1.1 through 3.1.5.

Section 3.2 focuses on modeling strategies and recommendations for development of new food safety process risk models in order to facilitate the application of sensitivity analysis methods.

3.1 Preparation of Existing Models for Sensitivity Analysis

The characteristics of existing models have a direct influence on the choice of sensitivity analysis methods and the scope of sensitivity analysis. In some cases, the modelers might not have anticipated the application of sensitivity analysis, and hence, the model may not have been developed in a manner that facilitates sensitivity analysis. To identify whether the modeling methodology used is compatible with application of sensitivity analysis, the model has to be

thoroughly reviewed and characterized. The following sections discuss the key features that need to be studied in the process of understanding a model.

3.1.1 Identification of Model Structure

The identification of model structure helps determine the scope of sensitivity analysis. Important model characteristics that are related to model structure include modularity, binning, and aggregation.

A general framework for performing quantitative food safety risk assessment is the modular process risk model (MPRM) (Nauta 2001a&b). Modularity is a way of organizing a model by breaking the overall model into sub-components. For example, each step or key activity in a food safety process risk model can be modeled in a separate module characterized by inputs and outputs. An example of modularity in food safety risk assessment is the *E. coli* model in which there are separate modules for processes such as slaughter of cattle and preparation of ground beef servings (FSIS, 2001). Conceptually, independent processes or components can be modeled as separate modules. Procedures that are repeated within the model also can be represented using modules. In a modular framework, an output from a module may be an input to other modules. In this manner, modules are connected to each other. For example, an output representing the contamination level of meat trim from the slaughter module can be an input to the preparation module representing the initial contamination of meat trim when modeling the grinding process.

The ability to apply sensitivity analysis typically depends upon a one-to-one correspondence between values of an output and values of an input. Thus, if the existing model utilizes a modular structure, in order to apply sensitivity analysis across modules in a framework composed of multiple modules, there should be one-for-one mapping of every input value to the output values, where the input of interest may be in one module and the output of interest could be in a dependent module. In a modular framework, modules can be connected in two ways: (1) serial connection; and (2) parallel connection. Figure 3-1 depicts a conceptual modeling framework comprised of three modules in series. Each module has exogenous inputs (i.e., $\overline{X}_1$, $\overline{X}_2$, and $\overline{X}_3$). Furthermore, modules that depend upon other modules also have internal inputs. The internal inputs are the outputs from the predecessor module (i.e., Y_1, and Y_2). Thus, a module output is influenced directly by internal and exogenous inputs for the immediately preceding module, and indirectly by the exogenous

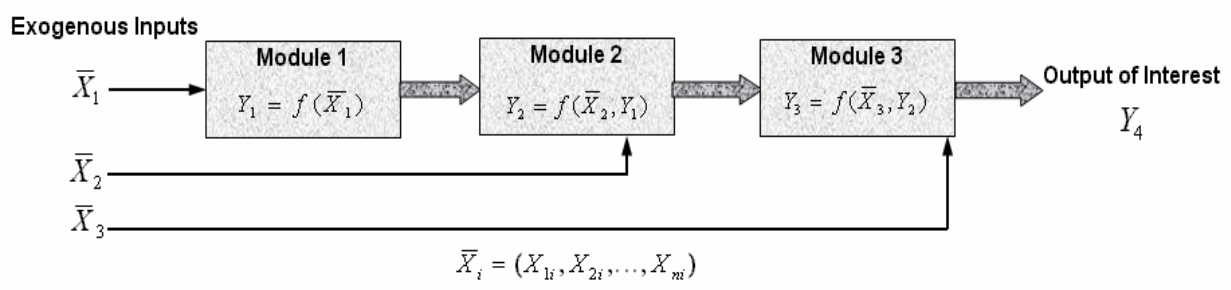

Figure 3-1. Conceptual Example of a Modular Modeling Framework Comprised of Multiple Modules in Series.

inputs for each predecessor module. Since the exogenous inputs typically include variability and uncertainty of interest to the analyst or decision maker, the typical desired objective of sensitivity analysis would be to evaluate the sensitivity of the selected output with respect to the exogenous inputs for all of the predecessor modules. A serial model structure that maintains the one-to-one correspondence between exogenous inputs and the selected output facilitates sensitivity analysis.

Modules can also be in parallel, as illustrated in Figure 3-2. The final model output is function of the outputs from each of several parallel modules. The key difference between the parallel and series frameworks is that in the parallel framework each module is independent from the other modules. This structure permits identification of the module to which the model output has the highest sensitivity on a relative basis. Frey *et al.* (2003) evaluated a case study with the *E. coli* model in which serving contamination, growth effect, and cooking effect modules had a parallel connection to the model output. A case study was prepared to prioritize the effect of these modules on the final exposure in ground beef serving. For the case study, each module was replaced by its corresponding output probability distribution. For evaluation of the priority rank of the distributions of outputs with respect to variability in exposure, four cases were examined using Monte Carlo simulation. Case *Zero* represented the situation in which all three variables (i.e., serving contamination, growth effect, and cooking effect) varied based on their distributions. For each of the other three cases, the output of one of the modules varied based on its distribution, while the other two outputs were conditioned at their mean values. Comparison of the results for these four cases helped identify which of the four cases had the largest maximum contamination or the highest probability of exceeding contamination levels that would be considered to be high. Results of the analysis indicated that cooking caused a maximum range of variation in the exposure to *E.*

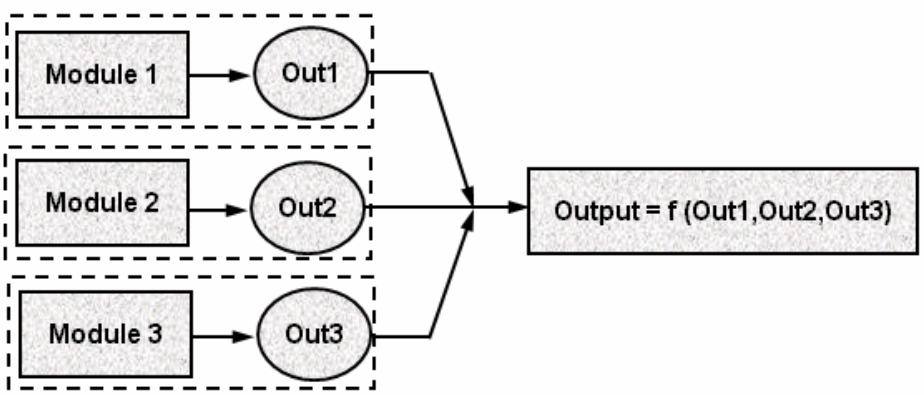

Figure 3-2. Conceptual Framework of Modules in Parallel.

Coli O157:H7 organisms and that growth resulted in the highest number of consumer exposures to *E. coli O157:H7* organisms in ground beef.

Sometimes modelers employ techniques that result in a loss of one-to-one correspondence between exogenous inputs to predecessor modules and the output of interest. These techniques might have been selected because they offered a computational convenience, provided a useful intermediate summary, or facilitated exchange of data between modules. Examples of these situations may include: (1) parameterizing empirical values of internal inputs into theoretical distributions; (2) summarizing continuous internal inputs into intervals (i.e., binning); and (3) aggregating which leads to many-to-one relationship between input and output values.

The one-to-one correspondence between an output and model inputs may be lost, for example, by using the inputs to estimate the parameters of a theoretical distribution that serves as an intermediate output for the model.

Another example leading to loss of one-to-one correspondence between an output of a module and its exogenous predecessor inputs is the use of binning for internal inputs. Binning is a commonly used summarization technique. For example, in the *E. coli* model the contaminant concentration in beef trimings is simulated as a continuous variable, but subsequently is binned by 0.5 log increments from 0 to 8 logs (FSIS, 2001). The estimated contamination for the combo bin is rounded to the next upper level. For instance, if the estimated contamination is 0.1 logs, it is binned into a range of 0 to 0.5 logs and is quantified as 0.5 logs. Another example where binning was used is the exposure module in *Listeria* model in which several individual meal servings were mapped to the same dose bin. After binning, the specific simulated meal serving

that resulted in a particular dose cannot be identified. The exposure range is binned into half logs. The fraction of serving contained in each bin is calculated (FDA, 2001).

Figure 3-3 illustrates how the binning approach leads to loss of one-to-one correspondence between an output and model inputs. The figure shows a continuous distribution of the pre-binned output values. In the binning step, n bins are defined using equal intervals. There is a one-to-one correspondence between the pre-binned output values and sampled values of the model inputs. For example, a value of y_j of the pre-binned output corresponds to a unique sample from the distributions of the model inputs. This figure shows that the problem arises when binned output values are used instead of the pre-binned output values. For example, if the j^{th} bin is selected from the binned output distribution, it is not possible to trace back the values of the inputs associated with the selected binned output value. This figure shows that there is a range of pre-binned output values associated with the j^{th} bin, and hence, there are not unique values of the model inputs associated with the estimated output value. In order to eliminate the binning in an existing model, it might be necessary to substantially change the model structure. Such a change may be beyond the scope of the analysis. Therefore, under this circumstance, modular-based sensitivity analysis may be considered instead.

Aggregation refers to situations in which multiple numerical values are combined into one numerical value. For example, in the *E. coli* model, the contamination level in a combo bin is estimated based upon contamination contributed by individual meat trims from multiple slaughtered cattle. Meat trim are from cattle that are slaughtered, dehided, and eviscerated, and are the result of a fabrication process. These steps take place in different parts of the slaughter plant. Each of these steps can be a source of contamination. The combo bin contamination level might be influenced by only a small proportion of the incoming meat trim. The combo bin contamination level is calculated based upon sum of the number of organisms from each of the possible sources of contamination. This type of situation is challenging to sensitivity analysis because there is a many-to-one correspondence between meat trim and the combo bin, whereas sensitivity analysis typically considers one-to-one correspondences. Sensitivity analysis could be based on one or more summary statistics. For example, the sensitivity of the combo bin contamination to the average meat trim contamination, or the most contaminated meat trim could

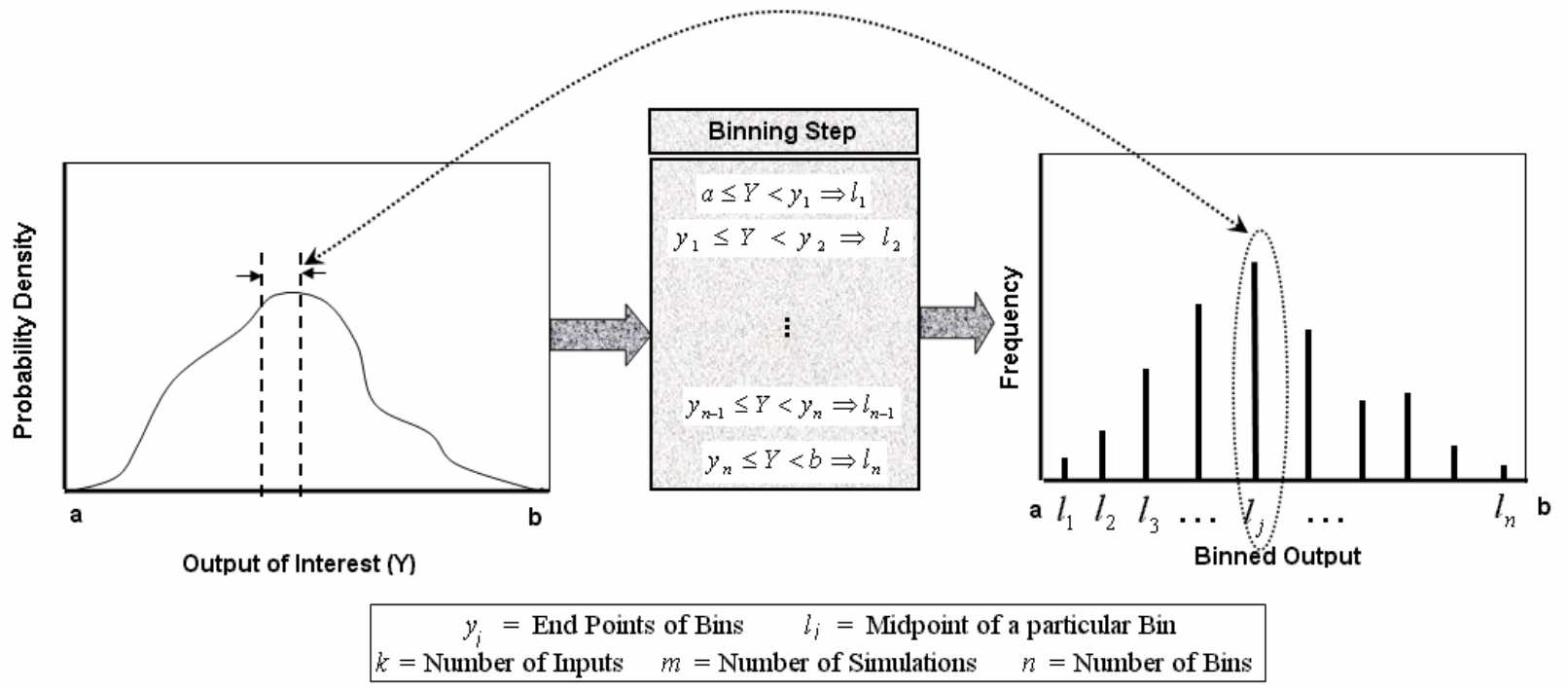

Figure 3-3. Conceptual Example of the Binning Approach, Leading to Loss of One-to-One Correspondence.

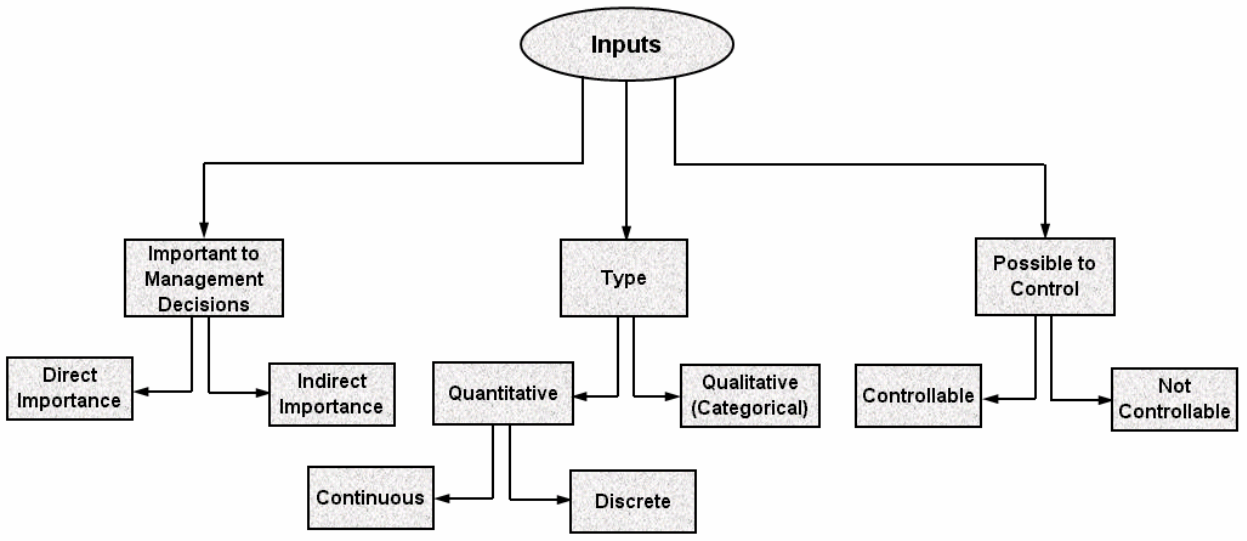

Figure 3-4. Possible Classifications of Inputs in Food Safety Process Risk Models.

be assessed. However, this would result in some loss of information. A simpler option is to perform sensitivity analysis only on internal inputs that depend on the model output of interest (e.g., the aggregated combo bin contamination levels), This would exclude consideration of inputs prior to the aggregation in the sensitivity analysis. For example, the original code in the *E. coli* slaughter module was modified to introduce a new composite animal that represents the characteristics of all contaminated cattle contributing to contamination in a combo bin. This composite animal has the total contamination from different contamination sources, and hence contributes the total number of *E. coli* organisms to a combo bin. Sensitivity analysis for combo bin contamination was performed with respect to the characteristics of the composite animal (Frey *et. al.*, 2003).

3.1.2 Identification of Inputs

To perform sensitivity analysis, the inputs of interest must be identified. Inputs can be classified in different ways. Figure 3-4 shows a possible classification of inputs in food safety process risk models. The inputs may be of direct or indirect importance to management decisions (e.g., serving size versus sanitation control). Some inputs may be controllable and others not (e.g., storage temperature versus lag period). Also, the degree of control may vary from one input to another. Inputs may be qualitative (categorical) (e.g., ground beef consumption types) or quantitative (e.g., storage temperature). Quantitative inputs can be continuous (e.g., storage time)

or discrete (e.g., number of infected animals in a truck). Some quantitative inputs may be described by empirical distributions, while other inputs may be represented by parametric distributions. The implications of the type of input with regard to the choice of sensitivity analysis method are discussed in Chapter 5 (see Table 5-1).

Model inputs may represent variability, uncertainty, or both in a probabilistic simulation. The manner in which uncertainty is modeled may vary among inputs. In some cases, uncertainty may be considered in the parameters of the variability distributions. For example, a triangular distribution may be used to represent variability in storage temperature. A triangular distribution is defined by three parameters representing minimum, most likely, and maximum values of the distribution. The most likely value of this distribution may be uncertain. A uniform distribution may be considered to represent the uncertainty for the most likely value. In contrast, uncertainty can be incorporated in an input as a choice of alternative variability distributions. For example, in a case of laboratory data for initial concentration of a pathogen, a few parametric distributions may be good fits to available data. There may exist uncertainty in selecting which parametric distribution better represents the data. Uncertainty can be addressed for the input by weighting fitted variability distributions and randomly selecting each variability distribution proportional to its weight during simulation. For example, four parametric distributions are considered for representation of variability in the initial concentration of the *Listeria monocytogenes* pathogens (FDA, 2001). These distributions include Beta, Weibull, Triangular, and Lognormal. Each parametric variability distribution is selected randomly in alternative realizations of uncertainty of the model to represent the uncertainty in the choice of a variability distribution.

3.1.3 Selection of a Model Output for Sensitivity Analysis

The selection of a model output for sensitivity analysis depends on the assessment objectives. For example, the output of interest may be different when objective varies from reduction in the risk of mortality in the population to reduction in the risk of mortality per serving. The relative importance of inputs is likely to vary according to the choice of output. An analyst should clearly determine the type of management objective that is expected before selecting the output of interest. Frey *et* al (2003) investigated case studies with the *E. coli* model in which the relative importance of inputs varied based upon the selection of output related to the risk reduction strategies. For example, when the number of contaminated ground beef servings was considered as the output of interest, the cooking effect and related inputs, including cooking

temperature and pre-cooking handlings, were identified as key. In contrast, when the amount of contamination in ground beef servings was selected as the output of interest, growth and related inputs, including storage times and temperatures, were characterized as key.

3.1.4 Simulation Design

Performing a simulation is a prior step to application of any sensitivity analysis method. Results obtained from sensitivity analysis are directly related to the characteristics and the scope of the simulation. An analyst should consider the following when designing a simulation: (1) efficiency; (2) coverage of rare events; and (3) burden-of-disease issues. These concepts are briefly discussed.

A simulation should be efficient. Apart from wasting computational resources, inefficient simulations also waste sensitivity analysis resources and undermine their ability to detect significant inputs. Many simulations are designed to provide exhaustive simulation of an event space even when the modeler is (or should be) fully aware that the level of risk in large portions of the event space is minimal in comparison to others. Examples of this include detailed simulation of temperatures below which there is certainly no growth, and detailed coverage of regions with very high lethality treatments where a disproportionate share of the risk is predictably concentrated in the regions of low lethality. Many of these situations would be greatly improved by partitioning the inputs to explore high-risk regions in more detail.

A well designed simulation should cover rare events. For most foods, the number of servings that result in illness will be a very small fraction of the total number of servings in the simulation space. The simulation can be partitioned to give greater coverage to these rare events. The implications of such a strategy for sensitivity analysis are that the simulation may be conditional on the probabilities associated with such rare events.

For many microbial pathogens, a large share of the burden of disease is borne by highly susceptible subpopulations. As a result, the simulation, and by extension the sensitivity analysis, should be focused on the important predictors of risk in these particular individuals. Thus, an implication is that the simulation should focus on subpopulations that are of special concern.

3.1.5 Probabilistic Simulation and Implications for Sensitivity Analysis

The scope and detail of inferences drawn from sensitivity analysis depend on the choice of probabilistic simulation dimensions. The presence or absence of variability and uncertainty dimensions in the modeling process should be identified prior to selection of a sensitivity

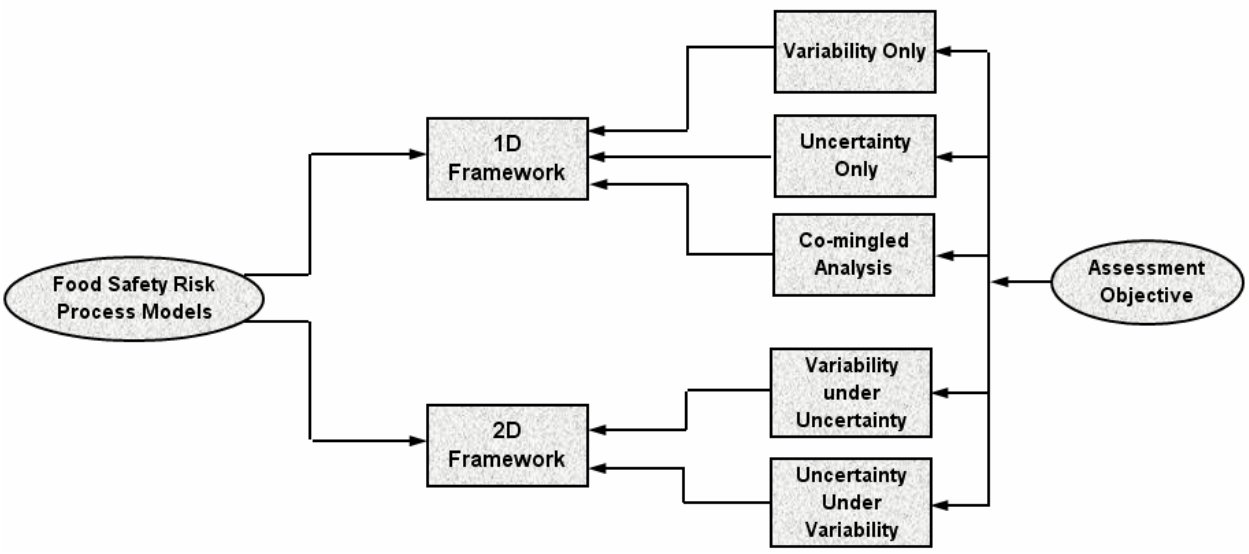

Figure 3-5. Alternative Frameworks for Food Safety Process Risk Models and Selection of the Probabilistic Dimension for Sensitivity Analysis Based on the Assessment Objective.

analysis approach. Food safety process risk models typically have one-dimensional or two-dimensional probabilistic frameworks incorporating variability, uncertainty, or both in the simulation. The selection of the specific probabilistic dimension of a model for sensitivity analysis depends on the assessment objective. Figure 3-5 shows a schematic diagram of alternative assessment objectives that are appropriate for the two modeling framework. The assessment objectives in this figure include: (1) variability only; (2) uncertainty only; (3) variability for different realizations of uncertainty; (4) uncertainty for different realizations of variability; and (5) co-mingled variability and uncertainty. The insights regarding sensitivity that can be obtained from these assessment objectives are discussed in Section 4.4.

For food safety process risk models that have a one-dimensional probabilistic framework, a practitioner may select variability only, uncertainty only, or co-mingled variability and uncertainty analysis. For example, Patil and Frey (2003) evaluated case studies with the FDA *Vibrio* (CFSAN, 2001) model that has a one-dimensional probabilistic framework considering only variability in the model inputs, and hence, sensitivity analysis was focused on variability. A two-dimensional framework may be used to quantify variability and uncertainty simultaneously. When analyzing the results, the analyst may prefer to focus on characterizing the range of uncertainty regarding estimated distributions for inter-individual variability in exposure and risk, or to focus on quantifying uncertainty in the exposure or risk for a particular fractile or

statistic from the variability dimension. In a two-dimensional framework, it is also possible to accommodate one-dimensional cases such as for variability only, uncertainty only, or co-mingled variability and uncertainty. For studies with major policy-based consequences, a practitioner may need to apply comprehensive sensitivity analyses that fully distinguish between variability and uncertainty depending on decision criterion. For example, when the decision criterion is to provide specific confidence that selected percentile in the exposed population falls below a threshold of concern, a practitioner would typically need to fully distinguish between variability and uncertainty. In contrast, for preliminary studies a practitioner may decide to focus only on variability or uncertainty.

Available resources also may affect the selection of the number of dimensions of the probabilistic simulation when performing sensitivity analysis. Two-dimensional sensitivity analysis is generally resource intensive. It demands computational resources both during the creation of datasets and the application of sensitivity analysis. The two dimensional analysis may not be practicable for sensitivity analysis methods that cannot be automated. In contrast, a one-dimensional analysis often requires fewer resources. Although application of sensitivity analysis on a single probabilistic dimension of the model is substantially easier, the credibility and robustness of the insights obtained with respect to the sensitivity may be reduced because the one-dimensional analysis often assumes that the other probabilistic dimension of the model can be replaced with deterministic values. For example, case studies performed by Frey *et al.* (2003) in which variability and uncertainty were fully differentiated show substantial ambiguity in ranks of key variability sources when uncertainty is considered. Quantifying the degree of ambiguity in the ranks of inputs helps in making management decisions that are robust to uncertainty.

3.1.6 Modification of Models

In some situations a model must be modified to apply sensitivity analysis. Some common modifications are discussed in this section. These modifications often demand changing data storage procedures, including model inputs, outputs, and internal inputs. Additional modifications may be required for some specific models.

3.1.6.1 Storage of Input and Output Values

In order to perform sensitivity analysis, a dataset including values generated during the probabilistic simulation for each input and corresponding output values should be available. As a common practice in programming, generated values of each input are stored dynamically in the

model. Typically, old values of an input from a previous iteration are replaced with new values to conserve the memory space and run time. Thus, it may be necessary to modify the model in order to store these values for future reference. Based upon the modeling environment, these values can be stored in a separate data file, spread sheet, or inside the model in an array. The software package used for application of sensitivity analysis may influence the format that would be used for data storage. For example, SAS© can read formats such as comma-separated values (.csv).

3.1.6.2 Storage of Intermediate Values

Storage of intermediate values estimated during the simulation of a model is useful in some cases. For example, in situations in which values of an internal input are binned before using in other parts of the model, it is helpful to store individual values of the intermediate input prior to binning. If the output of interest is calculated in the original model based upon a binned internal input, then the model should be modified, as discussed in Section 3.1.1, so that the output is calculated for individual values of the internal input.

3.2 Preparation of New Models for Sensitivity Analysis

To facilitate the application of sensitivity analysis when developing a new model, it is essential to consider the proper use of modeling techniques and model implementation. Good software engineering practices in general help application of sensitivity analysis (Frey *et al,* 2003). This section discusses some measures that can be used to facilitate sensitivity analysis when building a new model. These measures involve modeling environments, different approaches for characterizing variability and uncertainty, modeling strategies, and model documentation, which are discussed in Sections, 3.2.1 through 3.2.4, respectively.

3.2.1 Modeling Environments

Modeling environments involve: (1) the format of data storage; (2) the use of add-in software packages; (3) programming environments; and (4) aggregation and binning. More details about these issues are given in Sections 3.2.1.1 through 3.2.1.4, respectively.

3.2.1.1 Format of Data Storage

It is desirable to store data as pairs of inputs and outputs to facilitate the application of sensitivity analysis. The stored format should be compatible or easily convertible to the formats accepted by software packages commonly used for sensitivity analysis. Examples of such formats are comma separated values (.csv) and free text format (.txt).

Storing data as pairs of inputs and outputs may not be an optimal solution, especially for models with an extensive number of inputs and outputs or a large number of iterations during probabilistic simulations. As an alternative storing approach, application of relational databases is recommended. Using a database for storing the results of the analysis can facilitate the process of managing and retrieving the information in the form of outputs and corresponding inputs. A well-defined database can decrease the required storage space and increase the speed of retrieving data. An example of this strategy to store data for performing sensitivity analysis is the Environmental Protection Agency (EPA), Stochastic Human Exposure Dose Simulation, (SHEDS) Pesticide model (Zheng and Frey, 2003).

3.2.1.2 When and Where to Use Add-In Software Packages

"Add-in" is a term used, especially by Microsoft, for a software utility or other program that can be added to a primary program. Examples of add-in software packages are Crystal Ball™ or @Risk, which are applicable to Microsoft Excel. The use of add-in software packages helps reduce the amount of programming. However, the add-in software packages typically do not provide access to the code used to perform the various operations, and thus restrict the flexibility of the model with respect to possible modifications for sensitivity analysis. Performing the operations inside spread sheets using add-in software packages makes it accessible to those without knowledge of the programming language, but may make it difficult to audit the model. For complex and large models coded in Excel sheets, it becomes very difficult to understand the flow of the model and connection between different parts and modules inside the model.

3.2.1.3 Programming Environments

The choice of programming environments depends on the skill of the modeler, use of add-ins, and the scope of the analysis. For models that are extensive and that will be used for multiple analyses, a programming language environment and good software engineering practices are recommended. The choice of modeling environment should account for a trade-off, if any, between the skills of the analyst, resources, anticipated needs for future model refinements, and desired flexibility with regard to sensitivity analysis. The codes for most of the operations used in food safety process risk models, such as Monte Carlo simulation, are available as open source codes on the internet that can be downloaded free of charge. Open source codes of other available codes, such as R are also available that could be used for sensitivity analysis. The use of object-oriented or modular codes to perform various operations facilitates

modification and enhancement of models. Also, well-documented codes can be easier for a knowledgeable analyst to audit compared to the add-in software.

3.2.1.4 Handling Binning and Aggregation in the Model

If there is some rationale for application of binning in a model, it should be provided as an extra functionality. The user should be able to access the actual numerical values of an output before binning without having to modify the model itself. Hence, pre-binned values should be stored. Aggregation results in similar restrictions for application of sensitivity analysis as does binning. If aggregation is part of a real food safety process under study, such as the aggregation of meat trims to combo bins in slaughter plants, then the aggregation may be unavoidable. In some cases, aggregation is done because of lack of information about individual processes and insufficient understanding of what is exactly happening from a biological viewpoint. Hence, different processes in the continuum of bringing food from farm-to-table may be aggregated into a black box. In these cases it may be practical to refrain from aggregation by collecting more detailed data regarding the process under study.

3.2.2 Characterizing Variability and Uncertainty in the Probabilistic Simulation

The choice of probabilistic dimensionality depends on the assessment objective. If the objectives include characterizing inter-individual variability in exposure or risk, or identifying key sources of controllable variability in order to inform risk management decisions, a probabilistic simulation of variability should be included. If the objectives include quantification of uncertainty due to lack of knowledge, evaluation of the robustness of risk management options, or identification of key sources of uncertainty, then a probabilistic simulation of uncertainty should be included. If the assessment objective motivates inclusion of both variability and uncertainty, then a choice can be made between a one-dimensional simulation in which variability and uncertainty are co-mingled versus a two-dimensional simulation in which they are distinguished. The former may be interpreted as treating variability as a source of uncertainty about a representative member of a population, while the latter is typically preferred if the objective includes characterizing exposure or risk to different member of a population. Although the number of dimensions of probability simulation for a food safety risk assessment model depends on specific problems under study, ideally, both variability and uncertainty should be characterized since the explicit separation of variability and uncertainty helps analysts or decision-makers to understand how model outputs might improve.

When both variability and uncertainty dimensions are incorporated into a model, special attention should be paid to the modeling strategies to make the separation of variability and uncertainty practical. When distinguishing between variability and uncertainty, an analyst may be interested in quantifying the range of uncertainty regarding the population distribution of variability and/or key sources of uncertainty for specific fractiles of the variability distribution of the output. The former is easily addressed in a two-dimensional simulation framework in which there are multiple realizations of the distributions of inter-individual variability. The latter case requires that the set of random numbers associated with the variability dimension of each model input remain constant in multiple realizations of uncertainty. Section 4.4.4 provides further discussion regarding the latter case. An example of a food safety process risk model following this modeling strategy is the *Listeria* model in RTE foods (CFSAN, 2003). In this model, the set of random numbers generated for sampling from variability distributions of inputs are stored. At each uncertainty realization of the model, the model used the stored random numbers for sampling from variability distributions of inputs.

3.2.3 Modeling Strategies

A basic strategy in building a new model is to develop independent modules to facilitate sensitivity analysis. An independent module is one that does not share inputs or internal data with its predecessor or successor modules, and hence, the module is self-sufficient for calculation of its output. This modeling strategy is most easily achieved when using modules in parallel, as shown in Figure 3-2. For modules in series, as shown in Figure 3-1, as long as the one-to-one correspondence can be preserved, it is possible to attribute variation in the output to the inputs of one or more modules.

The main functionalities in a model should be implemented in separate modules. For example, there might be separate modules for random number generation, core model simulation, simulation result summarization and presentation, sensitivity analysis module, and others. An example of the use of such a strategy is the C++ version of the EPA/SHEDS/Pesticide model (Zheng and Frey, 2003). A modular framework can easily incorporate future additions of new functionalities simply by adding new modules without requiring extensive modification of the original model.

3.2.4 Model Documentation

Good documentation of the model structure and implementation is necessary to facilitate future application or modifications of a model. Good documentation includes:

- Introducing the model and its probabilistic framework
- Explaining of the model equations
- Introducing the model inputs and corresponding probability distributions
- Introducing the model outputs

It is necessary to introduce a model clearly and in detail, including mathematical equations used to describe the model and how the model was implemented in a computer program. Good documentation will also help the application of sensitivity analysis. For example, it will help an analyst to better understand a model including its inputs and outputs, and to select appropriate methods to perform sensitivity analysis. Good documentation requires that the model structure, the composition of model components, assumptions used, and software structure design implementing the model be clearly presented. For example, each module in a model should be clearly explained regarding its purpose, associated assumptions, its interior functional relationships, and its inputs and outputs. Each input should be clearly described regarding its type (e.g., probability distribution, numerical value or logical decision variables, and quantitative versus qualitative), values, and associated variability and/or uncertainty dimensions.

3.3 Summary

This chapter presented guidance on how to facilitate sensitivity analysis when working with an existing model or a new model, with a focus on food safety process risk models. For an existing model, in order to determine appropriate methods and scope for performing a sensitivity analysis, it is necessary to clearly identify model structure, model inputs, internal inputs used in models associated with binning and aggregation of model output of interest, and the type of probabilistic simulation. For a new model, recommendations that can facilitate sensitivity analysis are presented. These recommendations involve the modeling environment, design of the probabilistic simulation, modeling strategies, and documentation of models.

The next chapter provides guidance on how to define a case scenario in order to apply appropriate methods to perform sensitivity analysis.

4 DEFINING SENSITIVITY ANALYSIS SCENARIOS

The objective of this chapter is to provide guidance regarding how to define a case scenario as the basis for performing sensitivity analysis. A scenario is a set of assumptions about the nature of the problem to be analyzed (Cullen and Frey, 1999). A case study scenario for a food safety process risk model is comprised of a set of assumptions. These assumptions may be based upon recommendations from food safety experts, stakeholders, risk managers, or combinations of all three. The scenario should be relevant to the assessment objectives.

The definition of the most relevant or important scenarios is especially crucial in situations for which there are limitations of time and other resources with respect to performing sensitivity analysis. Thus, it is important to identify scenarios that are the highest priority for evaluation. A well defined case scenario will help concentrate the sensitivity analysis on the areas that are of more interest to risk managers and decision makers.

Figure 4-1 shows a schematic diagram of the case study scenario components for sensitivity analysis of food safety process risk models. This figure shows that issues such as the specific pathogen under study, susceptible subpopulations, pathways of exposure, temporal dimension, geographic extent of the analysis, and food categories considered in the analysis should be clearly specified in the case scenario. Furthermore, the probabilistic dimensions that are intended to be the focus of sensitivity analysis should be clearly specified in a case scenario.

In the following sections, the major components of a typical study scenario are explained. These components are classified into four categories including: (1) identification of the pathogens of interest and susceptible subpopulations; (2) identification of pathways of interest and selected food categories; (3) spatial and temporal dimensions of the model simulation; and (4) probabilistic features. Sections 4.1 through 4.4 provide discussions for these topics, respectively.

A scenario that does not fully address all of the attributes relevant to a particular study objective could lead to incorrect or incomplete insights regarding exposure and risk. For example, a scenario may not include all important food groups that might lead to high exposures to a particular pathogen. The failure to include a relevant attribute of the scenario is an error of omission. Errors of omission could either bias the total exposure and risk estimates, creating a

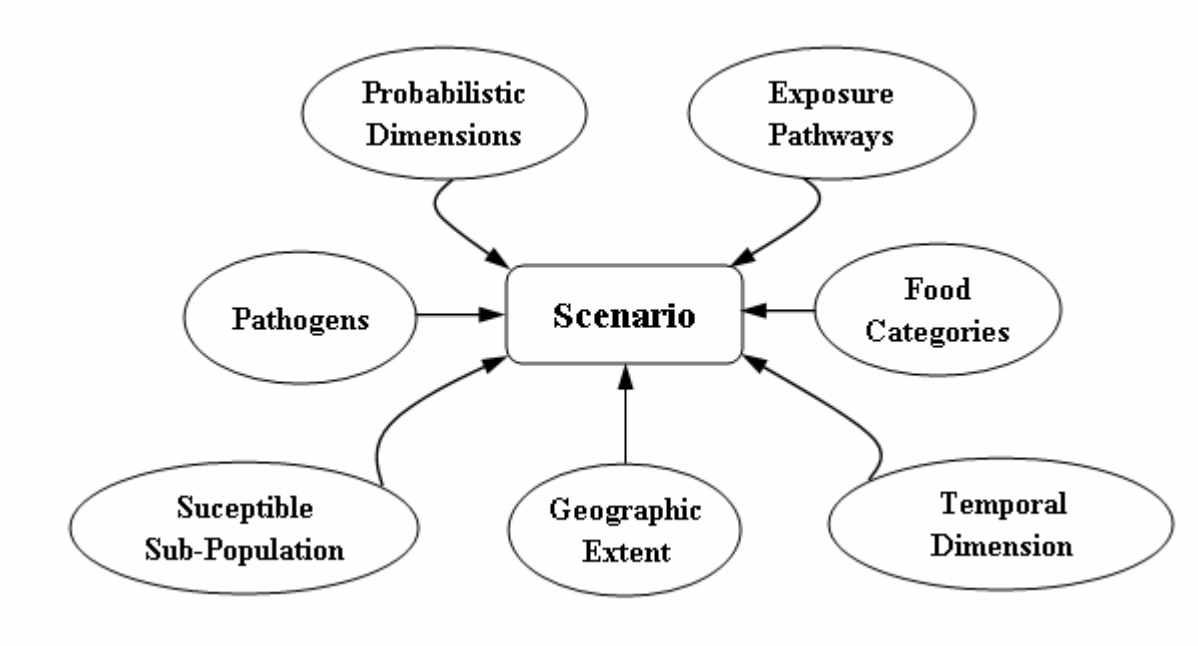

Figure 4-1. Components of a Scenario for a Microbial Pathogen Risk Assessment.

systematic error in the case scenario results, or for the cases that the model is calibrated incorrectly, an inability to attribute exposure to the correct pathways. The uncertainty associated with mis-specification of a scenario is discussed in Section 4.5. Section 4.6 provides a summary for the chapter.

4.1 Identification of Susceptible Subpopulation

In practice, many food safety risk assessments focus on a single pathogen. A pathogen of concern is typically identified because of a history of a high incidence of sporadic cases, known prevalence, or anticipation of future prevalence and possible outbreaks. Examples of assessments for individual pathogens include *Listeria monocytogenes* (CFSAN, 2003), *E. coli O157:H7* (FSIS, 2001), *Vibrio parahaemolytics* (CFSAN, 2001), and *Salmonella* (FSIS, 1998). A typical focus in these types of assessments is on acute illness associated with short term exposure to perhaps only one contaminated serving. The choice of pathogen will typically have implications regarding the selection of target subpopulations for exposure assessment, since only those subpopulations that consume a particular type(s) of food would be exposed.

The scope of the risk assessment is typically on specific population groups in which the risk of adverse effect due to exposure to the hazard is expected to be significant. For a particular pathogen, there may be multiple susceptible subpopulations because different groups of people

are exposed to a particular type of contaminated food. Susceptible subpopulations can be incorporated in the model using alternative dose-response relationships. Furthermore, exposed subpopulations could be differentiated by age classes with respect to food consumption patterns. Two examples are described here to illustrate how susceptible subpopulations can be addressed. One example is based upon the *Listeria* risk assessment model for selected ready-to-eat foods, and the other example is based upon the *E. coli O157:H7* process risk model for ground beef.

The *Listeria* risk assessment model considers three susceptible subpopulations based on age (CFSAN, 2003), including perinatal, elderly, and intermediate age. The perinatal group includes pregnancy-associated cases in which exposure occurs most often *in utero* as a result of foodborne *L. monocytogenes* infections of mothers during pregnancy. The elderly group includes individuals who are 60 or more years old. This group is thought to have higher susceptibility to *Listeriosis* due, in part, to physiological changes associated with the natural aging process. The intermediate age group is the remaining population, inclusive of healthy individuals and specific subpopulations highly susceptible to *Listeriosis*, such as AIDS patients or those taking drugs that suppress immune systems. Adjustment factors were used in order to incorporate available epidemiologic data into estimated dose-response functions specific to each group. Because of the availability of three groups in the model as part of the scenario definition, it is possible for an analyst to choose to focus on one group. For example, an analyst may decide to focus the scope of sensitivity analysis on perinatal subpopulation and identify key contributors to probability of risk of death or illness due to exposure to *L. monocytogenes*.

The *E. coli O157:H7* food safety process risk model for ground beef classifies the exposed population into four age categories. These age categories include: (1) 0 to 5; (2) 6 to 24; (3) 25 to 64; and (4) above 65 years of age. For each age category, different consumption data regarding the serving size in grams are provided in the model (FSIS, 2001). Although people in any age category are susceptible to infection with *E. coli O157:H7,* the greatest concern has most often been with respect to children under 5 years old and elderly people above 65. Thus, an analyst may decide to select these two subpopulations as the focus for sensitivity analysis.

If sensitivity analysis is focused on a subset of the subpopulations included in the scenario and the model, consideration should be given as to whether the insights regarding key critical control points, suggested critical limits, key sources of variability, or key sources of uncertainty are representative of other groups. If there is reason to believe that these insights

would differ among different subpopulations (e.g., because of differences in serving size, frequencies of servings, dose response, etc.) then it will typically be appropriate to perform sensitivity analysis separately for each subpopulation of most concern.

4.2 Identification of Pathways of Interest and Selected Food Categories

In order to adequately characterize total exposures, a scenario may need to consider multiple pathways of exposure. For food safety process risk models, the exposure pathway of most interest is typically ingestion associated with consumption of foods. Therefore, in defining a scenario, a choice must be made regarding which types of food to include.

As an example, the *Listeria monocytogenes* model includes 20 RTE categories. Not all of these categories contribute equally to exposure and risk. In fact, only a few contribute substantially to total exposure. Therefore, sensitivity analysis can be focused on a subset of the food categories.

The identification of priorities for scenarios involving food categories may be based upon a variety of factors. For example, hotdogs could be selected since it is the most important source of *Listeriosis* based upon survey data (CSFII, 1996 and DHHS, 1998). Milk might be of concern because it has a high consumption rate even though the prevalence of *Listeria monocytogenes* is low. The prevalence in smoked seafood is high. The largest outbreak of *Listeriosis* in the US was attributed to fresh soft cheese (CFSAN, 2003). These examples illustrate possible rationales for choosing to focus on a smaller set of food groups with regard to sensitivity analysis.

4.3 Identification of Spatial and Temporal Dimensions of Case Study

A risk assessment scenario includes spatial and temporal dimensions. The spatial considerations for a scenario could include the geographic area relevant to the farm-to-table continuum for each included food. Furthermore, the size and location of specific subpopulations may be of concern. Depending upon the level of model detail, it may be necessary to incorporate multiple scales of geographic information. For example, with regard to farming or processing plants, it may be necessary to consider local factors that could lead to contamination of food (e.g., infection for a particular feedlot of cattle). However, it also may be necessary to consider the throughput at a national or other scale in order to support estimation of the total number of exposures that lead to health effects in a population. There may be variability in the prevalence of foodborne pathogens with regard to geographic location because some regions may have more favorable conditions for growth and prevalence than others.

The temporal considerations for a scenario typically include: (1) the time for each major step in the farm-to-table continuum, in order to estimate the growth or inactivation of pathogens at each step; (2) the activity patterns of consumers with regard to frequency of consumption of particular types of foods; (3) "temporal dynamics" effects, whether at a short time scale (e.g., daily, weekly) or a longer scale (e.g., monthly, seasonal, annual); and (4) the time period associated with occurrence of illness as a result of one or more exposures. For example, pathogens may be more prevalent or more easily spread in one season than others. Implicit in a dose-response relationship is an averaging time over which an exposure produces a health effect. For example, a model may deal only with acute effects of the most recent exposure, whereas in other cases consideration might be given to cumulative exposures that result from multiple contaminated servings.

As an example of temporal considerations, the *E. coli* model considers the effect of seasons on the estimation of the *E. coli* organism prevalence within animals in a feedlot or a herd. Two alternative seasons are considered, including a high prevalence season (summer) and a low prevalence season (winter). The fact that prevalence of the *E. coli* organisms is higher in summer suggests that the case study scenario for sensitivity analysis in the model should focus on this time period.

4.4 Probabilistic Approaches

A scenario will typically include variability in exposures among different members of a population, unless the scenario is for a single individual. Regardless of whether the scenario is for a population or for a single individual, there will typically be uncertainty in the inputs to a model. Therefore, as a part of defining the scenario that is the basis for a risk assessment, it is important to define whether the assessment will explicitly incorporate variability, uncertainty, or both. Furthermore, there are alternative methods for dealing with variability and uncertainty. The choice of an appropriate method should be made taking into account the assessment objectives, the data quality objectives, the availability of data, and the importance of the assessment.

For purposes of providing guidance regarding the choice of the probabilistic component of a scenario, we consider five cases: (1) variability only; (2) uncertainty only; (3) variability for different realizations of uncertainty; (4) uncertainty for different realizations of variability; and (5) co-mingled analysis of variability and uncertainty. The applicability of each of these situations is briefly discussed.

4.4.1 Variability Only Analysis

The purpose of an analysis of variability only is to quantify inter-individual variability in exposure and risk. Variability only analysis also can be use to characterize a frequency or rate of occurrence in a group (e.g., the U.S. population) or series (e.g., a production line). Such an analysis is typically predicated on the assumption that the range of variability is much larger than the range of uncertainty; therefore, a judgment is made that uncertainty can be neglected. The appropriateness of this assumption will depend upon the specific problem and the objectives of the analysis. Variability can include controllable or explainable sources of variation (e.g., differences in refrigerator temperatures among a population of refrigerators) or stochastic sources of variability (e.g., differences in susceptibility to illness among individuals). In some cases, an analyst may prefer to stratify the analysis into each of several subgroups rather than perform one analysis in which distributions of variability are assigned to account for the presence of multiple subgroups.

For analysis of variability only, probability distributions that describe the variability in each model input are specified and can be propagated through the model using Monte Carlo simulation or similar techniques. Sensitivity analysis applied to the results of a probabilistic simulation of variability only can provide insight regarding the key contributors to variability in the model output of interest, such as exposure or risk. Information regarding key sources of variability that are controllable can be used to prioritize potential CCPs and to identify critical limits. Information regarding key sources of variability that are uncontrollable might imply priorities for data collection or research to improve the characterization of variability.

As a practical matter, it is unlikely that uncertainty is small relative to variability, at least for some model inputs. Therefore, the variability only approach may not be applicable in most situations.

4.4.2 Uncertainty Only Analysis

An analysis of only uncertainty may be appropriate if the range of uncertainty for each model input is large compared to the range of variability, or if the analysis is for a scenario in which the values of variable inputs are fixed. The latter might be the case, at least approximately, if the analysis is stratified. For example, an analysis could focus on one set of conditions in a scenario (e.g., one type of pathogen, one type of food, a specific serving size, a particular class of

individuals, etc.). An uncertainty analysis allows the use of incomplete information without producing a result that may appear to be misleadingly certain.

Typically, an uncertainty analysis is performed by specifying probability distributions for each uncertain input, propagating the distributions through a model using Monte Carlo simulation or related methods, and quantifying the uncertainty in a model output. Sensitivity analysis can be applied to the results of an uncertainty analysis in order to identify key sources of uncertainty. Knowledge of key sources of uncertainty can be used to prioritize additional data collection or research in order to improve the state of knowledge, and thereby reduce uncertainty. The results of an uncertainty only analysis with regard to exposure or risk are conditional on the fixed values of the variable inputs. If uncertainty is truly much larger in magnitude than inter-individual variability, then it will be difficult or impossible to make a distinction among individuals regarding exposures and risk.

The uncertainty only situation could arise in practice depending upon the assessment objectives or the state of knowledge (e.g., availability of data) pertaining to the assessment.

4.4.3 Variability Analysis for Different Uncertainty Realizations

Variability and uncertainty may both be of importance to an assessment. There are a variety of ways to account for both variability and uncertainty in a probabilistic analysis. In this section, the focus is on quantification of variability conditional on different randomly selected values from uncertain quantities. In the next section, the focus is on quantification of uncertainty for different fractiles of variability. Section 4.4.5 deals with an analytic approach in which both variability and uncertainty are included, but co-mingled. The reader interested in more detail regarding two dimensional approaches for dealing with variability and uncertainty is referred to Cullen and Frey (1999), Nauta (2000), and Vose (2000).

In variability analysis for different uncertainty realizations, the objective of analysis is to distinguish between variability and uncertainty. The focus of sensitivity analysis in this approach is to identify the key inputs for each realization of uncertainty. A realization refers to one model simulation based upon one randomly sampled value for each probabilistic input. Application of sensitivity analysis to this kind of case study provides insight regarding whether the identification of the key sources of variability is robust with respect to uncertainty. Figure 4-2 shows a schematic diagram of a case study scenario for application of sensitivity analysis to identify key sources of variability for different uncertainty realizations of a model. Each dataset

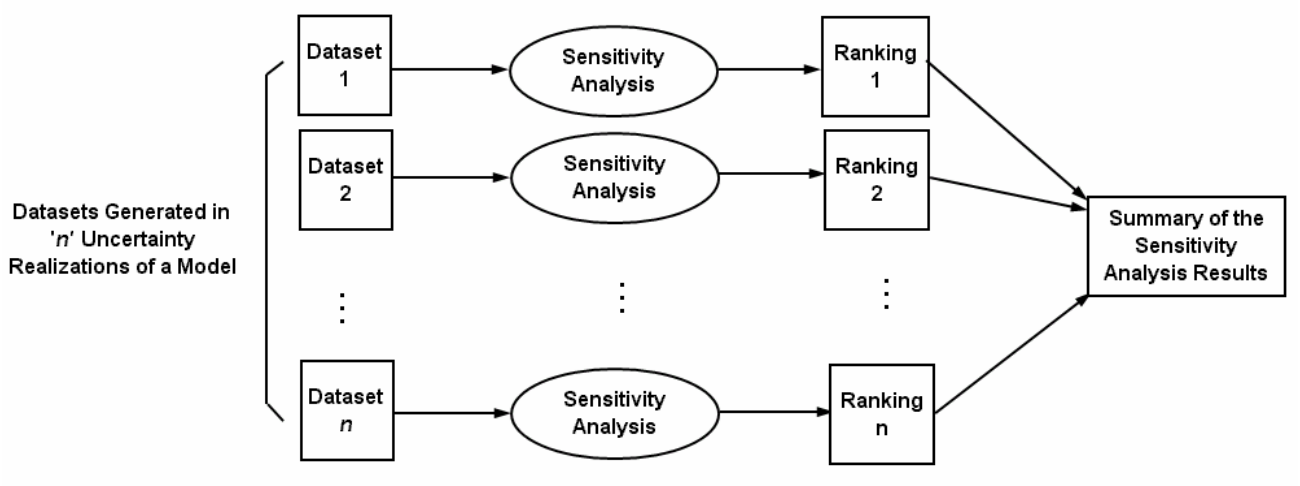

Figure 4-2. Case Study Scenario for Application of Sensitivity Analysis to Variability Analysis for Different Uncertainty Realizations of a Model.

in this figure includes randomly generated values from variability distributions of each model input for a given uncertainty realization and the corresponding model output values. In this case, sensitivity analysis is applied separately to each uncertainty realization. For each realization, the key sources of variability, critical limits, or both, are identified. This process is repeated *n* times to arrive at different sensitivity rankings for variability inputs in which *n* refers to the number of uncertainty realizations. The distribution of the rankings estimated based upon the *n* realizations can be used to assess whether an input is unambiguously important with respect to variability. Figure 4-3 shows an example of such a dataset for a model that has *k* inputs and the number of variability iterations is *m.* When performing two-dimensional Monte Carlo simulation, it is preferable to invest more iterations in the variability dimension than the uncertainty dimension, since the extreme tail of inter-individual variability in exposures is of interest.

To the extent that the sensitivity analyses yield similar results about the rank ordering of key inputs regardless of uncertainty, an analyst or decision maker will have greater confidence that the results of the analysis are robust to uncertainty. If the ranking of key inputs changes substantially from one realization of uncertainty to another, the identification of key inputs would be uncertain. Additional data collection or research may reduce this ambiguity.

Frey *et al.* (2003) evaluated a few case studies with the *E. coli* model featuring sensitivity analysis applied to variability for different uncertainty realizations. For example, this approach was applied to the growth estimation part in the *E. coli* model. Results of the analysis showed

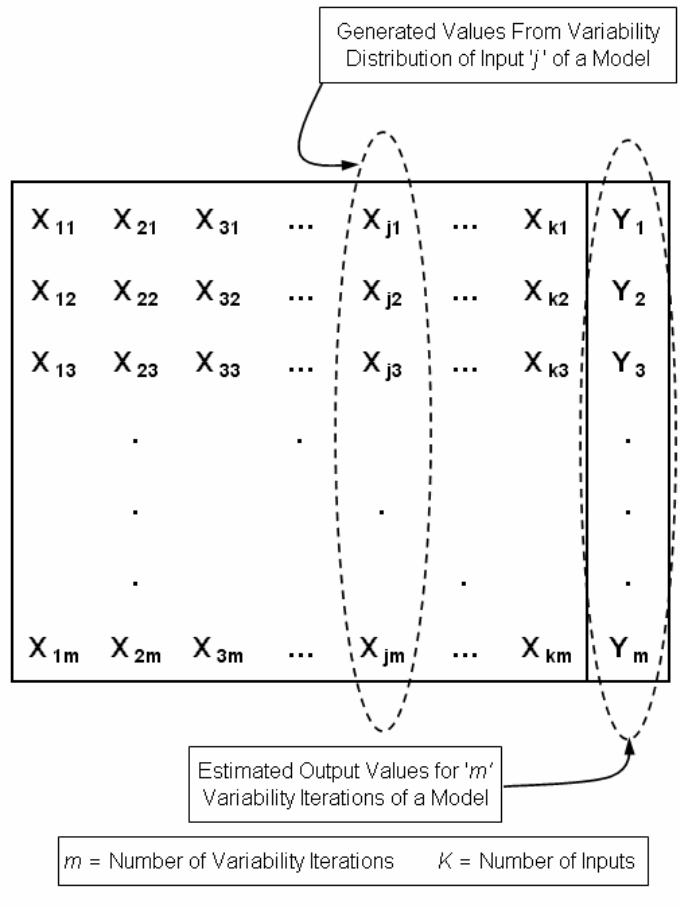

Figure 4-3. Dataset Including Generated Input Values from Variability Distributions and Estimated Output Values for a Specific Uncertainty Realization of a Model.

that selected inputs were consistently identified in the group of most important inputs in different uncertainty realizations of the model. Those inputs included storage times and temperatures at retail stores and homes. A few other inputs were identified that had wide ranges of ranking in different uncertainty realizations of the model. For example, the relative importance of storage temperature during the transportation step varied between a high ranking and ranking that was not statistically significant. The ambiguity in the true rank of storage temperature at transportation step was attributed to uncertainty in the input.

4.4.4 Uncertainty Analysis for Different Variability Iterations

The uncertainty analysis for different variability iterations is substantially similar to the probabilistic approach discussed in Section 4.3.3, except that this type of analysis focuses on the identification of key contributors to the uncertainty in the model output conditional on different

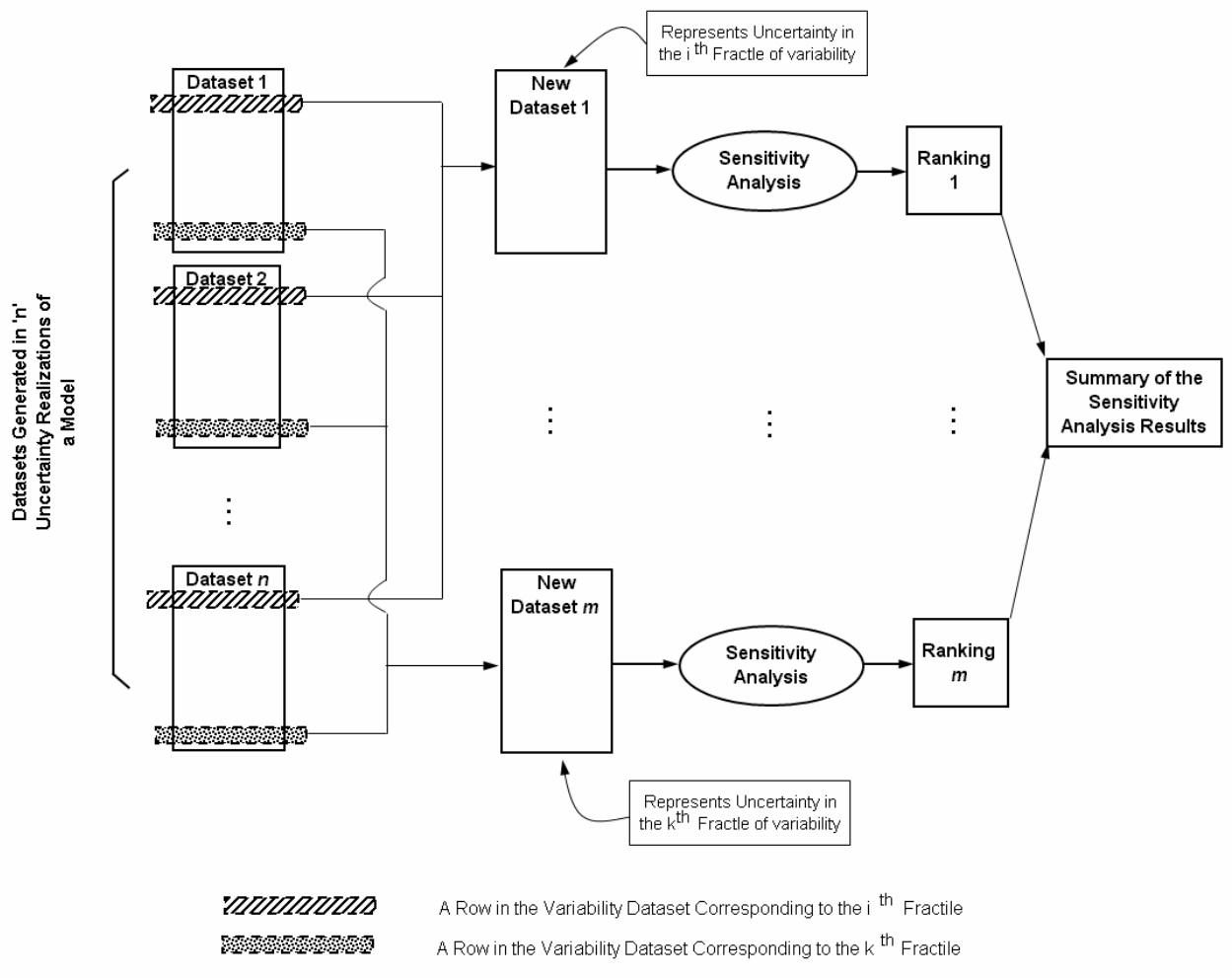

Figure 4-4. Case Study Scenario for Application of Sensitivity Analysis to Uncertainty Analysis for Different Variability Iterations of a Model.

values for quantities that are subject to variability. Results of this analysis will provide insight regarding how key sources of uncertainty for exposure or risk differ for different combinations of values for variable inputs.

Figure 4-4 shows a schematic diagram of a case study scenario for application of sensitivity analysis to uncertainty for different variability iterations. Datasets are shown on the far left for randomly generated values from variability distributions of each model input for each of *n* uncertainty realizations and for the corresponding model output values. Each dataset is similar to the example shown in Figure 4-3. In each dataset, there are *m* rows representing randomly sampled values from *m* fractiles of variability distributions. For a dataset, each row corresponds to a unique fractile of a variability distribution of model inputs. Prior to application

of sensitivity analysis, the *n* datasets should be reorganized into up to *m* new datasets. Each of the new datasets includes all of the uncertainty realizations corresponding to a specific fractile of the variability distributions. For example, the first new dataset shown in the figure includes *n* rows that hold randomly sampled values from the i^{th} fractile of the variability distributions and corresponding output values from each of the *n* original datasets. The variation between the values of rows in a new dataset is due to uncertainty in that specific fractile of the variability distributions.

In the next step, sensitivity analysis is applied to the new datasets and inputs are ranked based upon the sensitivity analysis results. Each set of rankings represents the key sources of uncertainty for a specific fractile of the variability. Comparison of the rankings for different variability fractiles can lead to key insights regarding sources of uncertainty. For example, the key sources of uncertainty in exposures for individuals who tend to have higher exposures may be different compared to those for individuals who tend to have lower exposures because of differences in the frequency or size of meal servings. If the ranking of the key sources of uncertainty substantially changes for different variability fractiles, decision-makers could focus on the key sources of uncertainty for the fractiles of most interest. For example, key sources of uncertainty can be identified for the upper 5 percent of the population exposed to the risk.

Application of sensitivity analysis to a case scenario using this type of probabilistic approach is informative for situations in which there is substantial uncertainty, especially regarding estimates of highly exposed individuals. A risk manager may prefer to make a choice of CCP or critical limits taking into account uncertainty for a particular portion of the most exposed subpopulation. However, if time and resources permit, knowledge of key sources of uncertainty for the most exposed or at risk portion of the population can be used to prioritize additional data collection or research that could reduce uncertainty. The assessment can be revised based upon new information, and a decision could be made at a later time based upon the reduced uncertainties.

4.4.5 Co-Mingled One Dimensional Variability and Uncertainty

There may be situations in which it is preferred to incorporate but not distinguish between variability and uncertainty. For example, during the process of model-building, the analyst may want to estimate the widest range of values that might be assigned to each model input for purposes of verifying the model and evaluating the robustness of the model to large

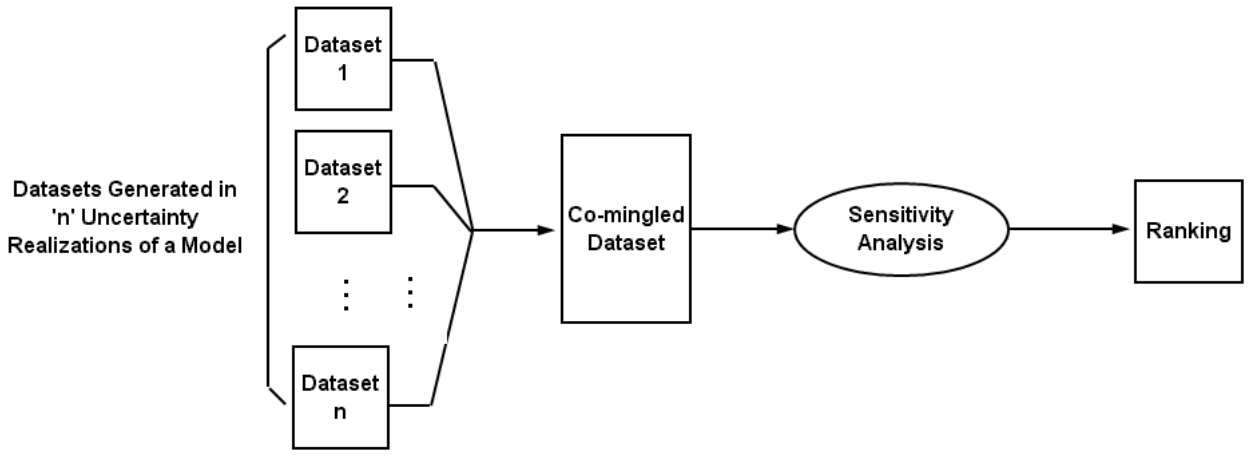

Figure 4-5. Case Study Scenario for Application of Sensitivity Analysis to the Co-mingled Variability and Uncertainty Analysis of a Model.

perturbations in its inputs. As a preliminary step in prioritizing data collection or the development of distributions for model inputs, an analyst may wish to assess the key sources of variation regardless of whether they represent variability or uncertainty. In some cases, an analyst may make a judgment that it is difficult to separate variability from uncertainty (e.g,. Nauta, 2000) or that a two-dimensional probabilistic approach is impractical or not necessary based upon the assessment objectives. However, if the assessment objective requires characterizing uncertainty associated with specific fractiles of the exposed population, then a two-dimensional approach is appropriate. Finally, if the assessment objective is focused on a randomly selected individual, rather than on specific fractiles of the exposed population, this type of probabilistic approach may be useful or appropriate to achieve the assessment objective.

Figure 4-5 shows a schematic diagram of a case scenario in which the probabilistic approach selected for sensitivity analysis is co-mingled variability and uncertainty. There are n datasets generated in n uncertainty realizations of the model. Each dataset is similar to the example shown in Figure 4-3, and hence, it includes m rows of randomly generated samples from variability distributions and the corresponding output values. Prior to application of sensitivity analysis, these datasets are appended together to form a single dataset with $n \times m$ rows. Sensitivity analysis is then applied to the co-mingled dataset and a set of ranking is obtained. By co-mingling variability and uncertainty in a single dimension, one typically would obtain wider ranges of values generated for model inputs during the model simulation compared to the situations in which only uncertainty or only variability had been characterized.

4.4.6 Selecting a Sample Size for the Probabilistic Approaches

In a probabilistic analysis that is implemented using a numerical method such as Monte Carlo simulation or related methods (e.g., Latin Hypercube sampling), it is necessary to specify a simulation sample size. The simulation sample size is the number of pseudo-random numbers that are generated for each model input according the probability distribution assigned to the input. The sample size represents the number of variability and/or uncertainty iterations used in a numerical simulation. Two constraints should be taken into account when selecting a sample size for a numerical simulation: (1) limitations of computer software, hardware, and time that impose an upper bound on the sample size; and (2) the acceptable degree of confidence or precision for model results, which impose a lower bound on the sample size (Cullen and Frey, 1999).

The sample size depends on the numerical simulation technique used in the model and the model output of interest. For example, typically a larger sample size is required when using Monte Carlo simulation as a sampling technique compared to Latin Hypercube sampling. If the central tendency of the output distribution is of more interest for analysts, a smaller sample size may be acceptable than if the focus is on the upper tail of distribution for a model output. For outcomes with a low probability of occurrence, it may be necessary to use large sample sizes or to stratify the analysis to focus only on situations that could lead to the outcome of concern. For example, there may be a very small probability that a food serving will be contaminated. If a small sample size is used, possibly no contaminated samples would be simulated. A large sample size may be required to obtain a sufficient number of simulated contaminated servings.

Alternatively, the analysis could be stratified to focus only on contaminated servings. In such a case, the results of a simulation would be conditional on the existence of contamination. A smaller sample size could be used since all samples would be contaminated. However, in order to assess the likelihood of exposure, the results of a conditional analysis would have to be multiplied by the probability of contamination.

4.5 Scenario Uncertainty

Because a scenario may fail to consider all of the factors and conditions contributing to variation in the output, uncertainty can be introduced. This source of uncertainty is known as scenario uncertainty (Cullen and Frey, 1999) and typically results in a bias. The sources of scenario uncertainty include descriptive errors, aggregation errors, errors in professional judgment, and incomplete specification of the scenarios (EPA, 1997).

Descriptive errors represent incorrect or insufficient information regarding a scenario. Sometimes, quantitative data regarding activity patterns leading to exposure is a convenience sample or represents only those producers who are required to file reports with the government under a particular regulation. For example, animal health surveillance systems should be updated using reports from farms in the region of study. Major farms that have substantial number of animals kept in feedlots or herds substantially contribute to the process of data collection. These farms are usually equipped with private veterinarians with sufficient laboratory equipment for better data collection. In contrast, small farms with a limited number of animals may not contribute to the surveillance system as they typically do not have regular sampling and testing procedures due to budgetary limitations. In the case that there is a large number of small farms in the region, descriptive error may be introduced into the scenario. In this case, the estimated infection rate in animals would fail to reflect animals from small producers that do not report adequate information.

Aggregation errors arise as a result of approximations or assumptions used in food safety risk models to simulate the process of bringing food products from the farm to the table. For example, assumptions regarding homogeneity of the consumer populations, temporal approximations such as assuming that the prevalence of the selected pathogen does not vary in different seasons, and spatial approximations such as assuming that the prevalence of the pathogen or consumption rate does not vary in different regions lead to scenario uncertainty in the analysis in the form of aggregation errors. Aggregation error can be considered as a source of systematic error in which bias may be introduced in the results of analysis.

Errors in professional judgment can affect every aspect of the exposure and risk assessment process. This type of scenario uncertainty may arise from failure to properly define exposure scenarios and select appropriate models for the analysis. Judgment errors also may be introduced because of limited knowledge of analysts. As an example, there may be a situation in which scientists assume that a pathogen under study can only be spread mechanically by the movement of animals, persons, vehicles and other pathways, and hence these pathways are incorporated in the model. Further research may reveal that airborne spread is a significant dispersal mechanism. Errors due to professional judgment introduce bias in the results.

Incomplete specification of a scenario is a source of uncertainty. For example, an analyst may neglect an important exposure pathway due to lack of information, or may fail to consider a

specific susceptible subpopulation in the analysis. For example, although an analyst knows that the chilling process in slaughter plants is a key step in contamination of carcasses with the *E. coli* organisms, lack of representative data and information of the real process of contamination in the chilling step may lead the analyst to neglect the simulation of this step in the model. Incomplete analysis will also introduce bias in the results.

The analyst should carefully review possible sources of scenario uncertainty specific to the assessment. All key assumptions that are the basis of the scenario should be clearly stated. Known descriptive errors and aggregation errors should be described. These errors may be unavoidable because of lack of data upon which to more fully describe categorical information or to disaggregate further. If possible, a qualitative or quantitative judgment regarding how much such errors may bias the assessment should be given. Errors in professional judgment would typically be uncovered if the work is exposed to review and comment. There may be professional disagreement regarding the importance of possible pathways not addressed in the assessment; if so, such disagreement should be acknowledged and a rationale for not including a particular pathway (or other possible component of an analysis) should be offered. The rationale could include a judgment regarding the magnitude of systematic or random errors associated with the omission. As new information becomes available, the key assumptions upon which the scenario is based should be revisited to determine whether the scenario should be revised.

4.6 Summary

This chapter provided suggestions on how to define a case scenario for performing sensitivity analysis. Key concepts were introduced and discussed as essential attributes of a scenario for a microbial risk assessment. An analyst should clearly define attributes such as susceptible subpopulations, geographic extent of the analysis, pathogen under study, exposure pathways, and the probabilistic simulation approach. Failure to sufficiently address these attributes may introduce error in the results of sensitivity analysis in the form of scenario uncertainty. Furthermore, the analyst should document the assumptions underlying the scenario and acknowledge possible sources of error.

The specification of a scenario can be done before a model is developed, in which case the model should incorporate all key elements of the scenario. If an existing model is to be used for a newly defined scenario, then consideration should be given to whether the model must be modified in order to accurately represent the desired scenario.

The setup of case study scenarios provides a basis for choosing appropriate sensitivity analysis methods. The next chapter introduces the available methods used to perform sensitivity analysis and provides guidance on how to choose appropriate methods based upon different situations.

5 SELECTION OF SENSITIVITY ANALYSIS METHODS

The objective of this chapter is to: (1) identify the key questions that should be addressed when selecting a sensitivity analysis method (Section 5.1); (2) provide a decision framework for making choices among alternative sensitivity analysis methods (Section 5.2); and (3) briefly review typical sensitivity analysis methods (Section 5.3). Section 5.3 focuses on the selected set of the most popular or accessible methods. Several additional methods that are promising but not yet widely used are discussed in more detail in Appendix A.

There are different ways of classifying sensitivity analysis methods. For example, these methods may be broadly classified as mathematical, statistical (or probabilistic), and graphical (Frey and Patil, 2002). Mathematical methods typically assess sensitivity of a model output to the range of variation of an input, and they typically involve calculating the output for a few values of an input within their possible ranges or for a small perturbation (e.g. Salehi *et al.*, 2000). Statistical methods involve running simulations in which inputs are assigned probability distributions, and the effect of variance in inputs on the output distribution is assessed (e.g. Andersson *et al.*, 2000). Graphical methods present sensitivity in the form of graphs, charts, or surfaces. Generally, graphical methods are used to give a visual indication of how an output is affected by variation in inputs (e.g., Geldermann and Rentz, 2001). We adopt this classification scheme as the basis for the brief review of methods in Section 5.3.

Alternatively, methods can be classified as screening, local, and global. Here, we explain the alternative classification scheme because it also provides a useful conceptual framework for comparing methods.

Screening methods are typically used to make a preliminary identification of the most sensitive model inputs. However, such methods are often relatively simple and may not be robust to key model characteristics such as nonlinearity, thresholds, and interactions.

Local sensitivity analysis focuses on relatively small perturbations near a fixed point in the model domain. For example, the sensitivity of various inputs for a particular fractile of the exposed population (perhaps representing a particular individual) could be the subject of a local sensitivity analysis. For small perturbations of the inputs, a linear approximation may be reasonable even if the model response over a larger variation of the inputs would be nonlinear.

Global sensitivity analysis methods must have the following two properties: (1) the sensitivity estimates of individual inputs take into account the effect of the range and the shape

of the probability distribution for each input; and (2) the sensitivity estimates of individual inputs are obtained while all inputs vary simultaneously (Saltelli, 2000).

This chapter is comprised of three sections. Section 5.1 presents a set of key questions for method selection. Answering those questions will help an analyst select an appropriate sensitivity analysis method for a particular case study. Section 5.2 presents procedures to guide the practitioner through the decision process of selecting a particular method. Section 5.3 provides brief background information for some of the most commonly used sensitivity analysis methods. Section 5.4 provides a summary for this chapter.

Key assumptions of this chapter are that the reader has already determined the reason for performing sensitivity analysis (Chapter 2), prepared the model for sensitivity analysis (Chapter 3), and defined the case study scenario (Chapter 4). Thus, there should be a clearly defined decision-relevant model output(s), a set of inputs of greatest interest, a well-defined scenario relevant to the assessment objective, and an appropriately structured model. Collectively, these issues comprise the problem definition for a particular analysis. The characteristics of the problem definition must be taken into account when selecting a sensitivity analysis method.

5.1 Key Questions for Selection of Sensitivity Analysis Methods

The selection of an appropriate sensitivity analysis method for a particular case study depends on the characteristics of the model and the case study. This section provides a series of eight key questions that the practitioner should consider before choosing a method. For each question, there is a brief discussion regarding the insight that an analyst may gain by addressing the question. These questions should be considered a starting point for selecting a method. The analyst should consider whether there are additional considerations in a particular case study beyond the issues addressed here. These questions are listed here and discussed in Sections 5.1.1 to 5.1.8.

- What are the objectives of sensitivity analysis?
- Based upon the objectives, what information is needed from sensitivity analysis?
- What are the characteristics of the model that constrain or indicate preference regarding method selection?
- How detailed is the analysis?
- What are the characteristics of the software that may constrain selection of methods?
- What are the specifications of the computing resources?

- Can "push-button" methods adequately address characteristics of interest in the analysis?
- Is the implementation of the selected sensitivity analysis method post-hoc?

5.1.1 What are the Objectives of the Sensitivity Analysis?

As the first step in selecting an appropriate method, an analyst should clearly define the objectives of sensitivity analysis with regard to the case study of interest. Knowledge of the objectives helps an analyst efficiently allocate the available time and resources to select an appropriate sensitivity analysis method that can provide relevant information and insight. Some common objectives of sensitivity analysis are: (1) rank ordering the importance of model inputs (e.g., critical control points); (2) identifying combination of input values that contribute to high exposure and/or risk scenarios; (3) identifying and prioritizing key sources of variability and uncertainty; (4) identifying critical limits; and (5) evaluating the validity of the model. Some of these objectives, such as (3) and (5), might primarily be research issues, while others are typically associated with risk management or regulatory decision-making.

Rank ordering the model inputs based on their contribution to the variation in the output of interest is often considered as a basic objective of sensitivity analysis of a food safety process risk model. An analyst can provide recommendations for model refinement through further data collection to improve model input distributions using the insight from rank ordering the inputs. As shown in the case studies explored by Frey *et al.* (2003), mathematical and statistical sensitivity analysis methods provide quantitative ranking of model inputs. Graphical sensitivity analysis techniques typically do not provide quantitative rankings of inputs.

An analyst may be interested in identifying ranges of key inputs that lead to high exposure and/or risk in food safety process risk models. For example, decision makers and regulatory agencies may be interested in identifying conditions responsible for high contamination levels in the process of bringing foods from the farm to the table. Some statistical methods, such as classification and regression tree (CART) and analysis of variance (ANOVA), can be used to identify combinations that lead to high exposures and risk (Frey *et al.*, 2003).

Objectives based upon identification and prioritization of key sources of variability and uncertainty in a probabilistic model are typically fulfilled using statistically-based sensitivity analysis methods. In addition, the analyst typically prefers that the sensitivity analysis method should provide quantitative measures of sensitivity for ranking key sources of variability and

uncertainty. Depending upon the assessment objectives with regard to whether and how variability and uncertainty should be distinguished, sensitivity analysis can be applied in a one-dimensional or two-dimensional probabilistic simulation framework. For example, if the assessment objective includes distinguishing between inter-individual variability in exposure versus uncertainty, then a two-dimensional probabilistic simulation framework would typically be employed. In such cases, sensitivity analysis can be applied to identify key sources of variability subject to multiple realizations of uncertainty, or vice versa. Ideally, in such situations the process of sensitivity analysis should be automated because of the potentially large number of model iterations required. As an example, Frey *et al* (2003) demonstrated the capabilities of regression analysis, correlation analysis, and ANOVA to identify key sources of variability and uncertainty when applied to two-dimensional food safety risk assessment models.

If the objective is to identify critical limits for controllable inputs, it is often necessary to use relatively sophisticated sensitivity analysis methods. This is particularly the case if the model has substantial nonlinearity, interactions, or thresholds. These methods often involve a probabilistic simulation (e.g., Monte Carlo simulation). Critical limits are sometimes associated with saturation points or thresholds in the input domain. A saturation point is a value for a model input above which there is no change in the output, but below which there can be substantial variation in the output. For example, there is an intrinsic upper limit on microbial pathogen growth rate. In contrast, a threshold is a value of a model input below which there is no change in the output. For example, if temperature is low enough, then the growth of a microbial pathogen may be essentially zero. Sensitivity analysis methods based upon linearity assumptions are not well-suited for identification of these kinds of limits. However, some statistical (e.g., ANOVA, CART) and graphical methods (e.g., scatter plots) can provide useful insights in these situations (Frey *et al*., 2003).

If the objective is to evaluate the validity of the model, then the analyst is typically interested in: (1) qualitative characteristics of how the output responds to changes in one or more inputs; and/or (2) how the output varies quantitatively as a result of varying inputs. Qualitative characteristic can include: (a) does the output increase or decrease if an input increases?; (b) does the output vary in a linear, monotonic, or non-monotonic manner; (c) are there threshold or saturation effects, or other characteristics of specific interest (e.g., inflection points, singularity points); and (d) are there interactions between two or more inputs that vary

simultaneously?. With regard to each of these characteristics, or perhaps others, the analyst will typically assess whether the model output is responding in an appropriate or explainable manner. If so, then confidence in the model tends to increase. If not, then corrective measures with respect to the model structure and assumptions are indicated. Sources of errors in a model response could include incorrect specification of the inputs or a problem with the analytical model or the computer code.

Quantitative characteristics are also useful to evaluate. For example, a key question is whether the range of the model output is appropriate with respect to variation of particular input. Perhaps the model seems to be far too sensitive to an input that was previously thought to be unimportant. Is this because of a mistaken prior perception, or did an error in the input data or the coded model produce an incorrect result? Does the model seem to be producing values far too high or low compared to what was expected? Perhaps there is a units conversion problem, data entry mistake, errant constant or parameter in the model or other mistake (e.g., incorrectly formulated or coded equation). The process of doing sensitivity analysis often helps the analyst gain critical insights into problems with the model or input data. Thus, sensitivity analysis should be done as part of model development to allow sufficient time to identify and correct possible problems.

5.1.2 Based Upon the Objectives, What Information is Needed from Sensitivity Analysis?

The analyst should decide prior to analysis as to what information is of most interest. Examples of information that could be useful, depending upon the objectives, include the following:

- Qualitative or quantitative ranking of inputs
- Discrimination of the importance among different inputs
- Grouping of inputs that are of comparable importance
- Identification of inputs that are not important
- Identification of critical limits
- Identification of inputs and ranges that produce high exposure or risk
- Identification of trends in the model response

Graphical sensitivity analysis methods typically provide qualitative insights regarding the importance of model inputs, whereas mathematical and statistical methods typically produce

quantitative measures of the sensitivity of an input. The analyst will typically want to be able to distinguish whether an input is of high, moderate, low, or no importance with regard to a particular output. Therefore, it is often useful to be able to discriminate regarding whether one input is more important than another input, even if the output is sensitive to both. Some statistical methods provide measures of whether one input is significantly more sensitive than another input. It is sometimes the case that several inputs will have comparable importance with regard to a particular output. The analyst may wish to group inputs that are of comparable importance based upon similar sensitivity index values. Similarly, the analyst may wish to identify inputs that are not important to the assessment. Information regarding unimportant inputs is useful because it is typically not necessary to devote additional resources to refining the assumptions for such inputs. Such resources can be prioritized to other inputs.

In order to identify critical limits in a model, the selected sensitivity analysis method should have the capability of comparing the model responses for different intervals for one or more model inputs. Similarly, if the objective is to identify high exposures, the selected sensitivity analysis method should identify the inputs and their ranges that are responsible for such exposures. Methods such as CART or ANOVA have capabilities to provide these kinds of insights.

Many sensitivity analysis methods provide information regarding trends in the model output, such as linear versus non-linear, ranges of variation, existence of thresholds, and other characteristics (e.g., does the output increase or decrease with respect to increase in a particular input). Graphical, mathematical, and statistical methods can provide a varying range of these kinds of insights.

5.1.3 What are the Characteristics of the Model that Constrain or Indicate Preference Regarding Method Selection?

Food safety process risk models typically have specific characteristics that may constrain the application of sensitivity analysis methods. These characteristics include: (1) nonlinearities; (2) interactions; (3) thresholds and saturation points; and (4) categorical inputs (Frey, 2002).

An ideal sensitivity analysis method should be model independent. Specifically, a sensitivity analysis method should not require any assumptions regarding the functional form of the risk model and should be applicable to different model formulations. Some methods are considered to be global and model-independent. An example is CART. However, some methods, including many commonly used methods, are based upon assumptions regarding the functional

form of the model. For example, sample correlation analysis or linear regression are based upon the assumption of a linear model. However, the model might not actually be nonlinear. In some cases, a linearizing transformation could be made, such as a log transformation, in order to have a more nearly linear model response to changes in inputs. Thus, with appropriate transformations, sometimes it is possible to apply these techniques. Rank correlation coefficients or rank regression are based upon the assumption of a monotonic model. If a sensitivity analysis method based upon an assumed functional form of a model is applied to a model with different characteristics, then the results of the sensitivity analysis may not be valid. For example, if method based upon linearity, such as nominal range sensitivity analysis (NRSA) or sample correlation analysis, is applied to a nonlinear or non-monotonic model, then the insights regarding sensitivity could be inaccurate or invalid.

Because interaction between inputs is often one of the common characteristics of food safety process risk models, an analyst should typically select a sensitivity analysis method that can deal with interactions. The most commonly used mathematical sensitivity analysis methods (e.g., NRSA, differential sensitivity analysis (DSA)) are not able to capture any interaction effects between model inputs. However, many statistical methods, such as ANOVA, CART, or appropriately specified regression approaches, can identify interactions between model inputs. Graphical methods may provide insights regarding interaction effects; however, the use of graphical methods typically requires expert judgment based upon visual inference and therefore is a qualitative approach to sensitivity analysis.

Identification of thresholds and saturation points that have substantial effect on the model output (e.g., exposure or risk) is useful because they may imply critical limits. Sensitivity analysis methods that enable an analyst to compare variation of the model output in different regions of an input domain are appropriate for identification of these model characteristics. Such methods include ANOVA, CART and some graphical sensitivity analysis methods. ANOVA provides direct statistical measures for tracking thresholds and saturation points. The interpretation of the results from CART is typically based upon visualization of regression trees in order to identify critical limits. Graphical sensitivity analysis methods provide visual insight regarding possible thresholds or saturation points in an input domain. For example, scatter plots can assist in identification of thresholds and saturation points taking into account simultaneous variation of other inputs. Graphical analysis based upon conditional sensitivity analysis can also

help identify thresholds and saturation points for an input conditional on point values from other inputs.

Categorical or qualitative inputs are commonly used in food safety risk assessment models. Examples include food categories, season, consumer age, eating location, and consumption type. Mathematical methods such as NRSA and DSA do not accomodate these types of inputs. However, an analyst could simply run the model with different values of such an input while other inputs are held constant and assess the range of variation in the model output compared to that obtained when each of several other inputs are varied individually. Among statistical sensitivity analysis methods, ANOVA and CART can deal with categorical inputs. Regression analysis can deal with categorical inputs by using dummy variables (Neter *et. al.,*1996, Frey *el. al.,* 2003). Graphical sensitivity analysis methods can be used to visualize the variation of model output versus different categorical values of inputs.

5.1.4 How Detailed is the Analysis?

Screening or refined sensitivity analyses can be conducted based upon the objectives of an analysis. The choice between the two is typically governed by resource availability, the importance of the analysis, and the stage of the assessment. Screening analyses require less resource than refined analyses. A screening analysis might be used in the early stages of model development to help refine the model and its inputs and to assess model validity. A refined analysis might typically be used in the later stages of analyses with a model that has undergone previous screening analyses. A refined analysis is preferred if the results will be used to make decisions regarding commitments of large amounts of resources for further model development, data collection for model inputs, or risk management strategies. Screening and refined methods can be applied as appropriate to one- or two-dimensional probabilistic simulations, although typically additional refinement of the sensitivity analysis would be accompanied by refinement in the risk analysis.

Screening and refined analysis can be used together. The time and effort to execute a refined analysis often depends on the number of inputs that are included in the analysis. Therefore, it is often useful to use a screening method to identify model inputs that are not important with regard to variation in the output of interest. The refined analysis can then be applied to a smaller set of inputs for which there is reason to believe that the model output has at least some sensitivity.

Typical screening analysis methods include local mathematical methods such as DSA or NRSA, as well as so-called "push-button" techniques that are readily available in commercial software packages. Examples of the latter include sample or rank correlation coefficients available in Microsoft Excel add-ins such as Crystal Ball™. The "push-button" methods for sensitivity analysis are further discussed in Question 7.

From case studies in Frey *et al*. (2003), it appears to be the case that if an input is identified as unimportant using a screening method, in practice it is often also identified as unimportant when using a refined method. When applying a screening analysis method, the analyst should keep in mind that there are several potentially critical limitations of such methods, including: (1) screening methods are often not model-independent, and therefore the insights may be inaccurate if the model structure deviates significantly from the theoretical basis of the method; (2) some screening methods are local, and cannot adequately account for simultaneous variation in multiple inputs; and (3) for local screening methods, it may not be possible to identify the inputs and ranges that lead to the highest risk or to gain useful insight regarding critical control points. Thus, it is preferable to confirm results from a screening method with a different method or to follow-up with a refined analysis.

5.1.5 What are the Characteristics of the Software that may Constrain Selection of Methods?

As a practical matter, analysts do not typically have the resources to develop computer code to implement sensitivity analysis methods and, therefore, will typically rely on readily-available software packages for this purpose. If the sensitivity analysis is to be conducted with an existing risk model, then the analyst may be forced to accept the constraints imposed by the existing software environment. In contrast, if the analyst is developing a new risk model, then the choice of programming environment should be made taking into account the anticipated needs with regard to sensitivity analysis, as discussed in Chapter 3.

For probabilistic models that are implemented in a spreadsheet environment such as MS Excel, the probabilistic capability may be available via a commercial add-in such as Crystal Ball™ or @Risk™, or via macros that have been developed by the analyst or others. Built-in sensitivity analysis methods, such as sample or rank correlation coefficients, are often included in these types of packages and are referred to here as "push-button" methods. Typically, there is a limited degree of automation of such methods. For example, it is usually possible to apply

push-button methods in a one-dimensional probabilistic simulation but it may be more challenging to apply them in separate dimensions of a two-dimensional probabilistic simulation.

In order to apply other sensitivity analysis methods, and especially refined methods, if a spreadsheet-based risk model is being used, it will typically be necessary to export to another software package sample values for the model inputs and outputs. For example, if the analyst wishes to make extensive use of regression methods, or to use methods such as ANOVA or CART, then it may be preferred to perform the analysis in specialized software such as SAS©, S™, S-PLUS™, or R. The first three are examples of commercial packages. Such packages can be expensive in terms of licensing fees as well as with regard to training and skill level requirements on the part of the user. The latter is a publicly available statistical package. In addition, there are software packages such as SIMLAB that include a variety of sensitivity analysis methods (SIMLAB, 2000).

If a new risk model is being developed, then the model can be designed to include sensitivity analysis methods as an integral part of the model or to export data in the format required by other software packages so that sensitivity analysis can be performed. For a highly refined analysis, the ideal situation is to include the preferred sensitivity analysis methods as features of the risk model to facilitate automated application of the method.

5.1.6 What are the Specifications of the Computing Resources?

The computing resources required for a given sensitivity analysis method are usually proportional to the resources necessary to perform a typical simulation with the model but also depend on the characteristics of the method. For mathematical methods, such as DSA and NRSA, it will typically be necessary to perform a number of model runs proportional to the number of inputs to be included in the sensitivity analysis. For statistical-based methods, sensitivity analysis can be performed based upon a typical probabilistic simulation. In this case, the number of model runs is determined by the data quality objective of the probabilistic simulation (e.g., the precision with which statistics of the model output should be simulated) taking into account the data quality of the model inputs. Computing constraints can include clock time, CPU time, data storage, other hardware issues, and software limitations.

Clock time refers to the total elapsed time from the start to finish of an analysis. CPU time refers to the actual amount of processor time required to conduct a simulation. Since most

risk models are implemented on desktop personal computers, the clock time is typically of more concern than the CPU time.

Data storage is a design consideration when selecting or developing a model, or when choosing a hardware platform upon which to run a model. Especially for two-dimensional probabilistic simulations, the analyst may have to be selective regarding which data to save in order not to overwhelm available data storage and handling capabilities. Thus, as described earlier, it can be helpful if a less resource intensive screening method can be used to focus a refined analysis on a limited set of model inputs.

Depending upon how the risk model is designed, it may be possible to allocate the computing task to several computers operating in parallel, such as in a distributed network. The use of such networks has been demonstrated for Monte Carlo simulations of a variety of model types, and leads to a reduction in clock time necessary to complete a probabilistic simulation. The ability to distribute the computing task among multiple computers will depend on the capabilities of the available network and availability of software to manage the parallel computing tasks. Depending upon the scope and importance of the analysis, it may be worthwhile to involve computer experts in the design and operation of a distributed network.

5.1.7 Can "Push-Button" Methods Adequately Address the Characteristics of Interest in the Analysis?

As described in previous sections, some sensitivity analysis methods are built-in features of commonly used software tools such as @Risk™ and Crystal Ball™ and are referred to here as "push-button" methods. Under pressure of deadlines and faced with constraints of an existing model, an analyst may prefer to use these methods rather than attempt to implement more refined techniques. However, a key question is whether such methods are adequate for the intended case study.

Push-button methods typically include sample (Pearson) and rank (Spearman) correlation coefficients applied in the context of a one-dimensional probabilistic simulation. These methods are applicable to global sensitivity of linear or monotonic models, respectively. Sample correlations are not a robust method for dealing with thresholds or interactions and may not provide useful insights when applied to categorical inputs. Rank correlations are more robust than sample correlations with regard to thresholds but are not capable of providing insight regarding the existence of a threshold or the value of the threshold. If an analyst chooses to proceed with a push-button method whose theoretical basis differs from the key characteristic of

the model to which it is applied, then the accuracy and robustness of the results cannot be assured.

In the long-run, it will be helpful to the food safety risk community if add-ins can be developed, either as shareware or commercially, that incorporate sensitivity analysis methods that are more compatible with the characteristics of the models.

5.1.8 Is the Implementation of the Selected Sensitivity Analysis Method Post Hoc?

A sensitivity analysis method is referred to as post hoc if it is applied to previously prepared results from probabilistic or deterministic simulation of a model but is not included as a component of the model itself. Therefore, the application of a post hoc sensitivity analysis method does not contribute to the process of model simulation, but it may impose requirements regarding the type and format of data that should be stored from the simulation. For example, regression analysis can be applied to a dataset of randomly generated values of the inputs and corresponding output values. In these cases, it is possible for the sensitivity analysis to be performed in a different software environment and by different personnel compared to the simulation modeling. Examples of methods that can be applied post hoc include sample and rank correlations, regression analysis, ANOVA, CART, and graphical techniques.

In contrast, some methods require a different simulation strategy and therefore may have to be programmed differently or implemented manually or interactively. Some of this will depend upon how the model is structured. For instance, to use NRSA, an analyst may need to change the values of an input to extreme values while holding the other inputs at their point estimates. However, if a model has been structured so that the process of varying one input at a time can be automated, then it could be possible to generate a dataset necessary for NRSA automatically and to analyze the data set post hoc.

Among statistical methods for sensitivity analysis, Sobol's method, Fourier amplitude sensitivity test (FAST), and mutual information index (MII) typically are not post hoc. These methods require approaches to model simulation different than the conventional Monte Carlo methods and therefore must be incorporated into the simulation process. The application of these kinds of methods typically requires more advance planning and coding than do the post hoc methods. Although these are powerful methods that offer advantages over push-button methods, their widespread practical application is limited until software becomes available by which these methods can be easily incorporated into a risk model. Alternatively, a risk model can be

incorporated into a sensitivity analysis framework such as SIMLAB in order to make use of some these methods. The recoding of a risk model into SIMLAB may involve trade-offs between SIMLAB compatibility and desirable features of the model (e.g., Patil and Frey, 2003).

5.2 Decision Framework to Assist in Selecting Sensitivity Analysis Methods

The goal of this section is to provide practical guidance regarding how to select a specific sensitivity analysis method depending upon the level of analysis to be done (i.e. screening versus refined) and the objective of the sensitivity analysis.

This section introduces two decision trees summarizing the discussions presented in Section 5.1 for selecting an appropriate sensitivity analysis method. The decision frameworks are shown in Figures 5-1 and 5-2. The sensitivity analysis methods used in these frameworks are briefly explained in Section 5.3.

Figure 5-1 shows that the first step in selecting an appropriate sensitivity analysis method is to decide the level of detail expected from sensitivity analysis. This figure presents two levels of sensitivity analysis: (1) screening analysis; and (2) refined analysis. For screening analysis, the choice of sensitivity analysis method depends on the simulation approach of a model. If a deterministic approach is selected for screening analysis, particularly for local sensitivity analysis, then mathematical sensitivity analysis methods are recommended. For probabilistic approaches, push-button techniques including sample and rank correlation coefficients are listed as appropriate methods. In contrast, if a practitioner decides to perform a refined analysis, the choice of a method depends on the objective of sensitivity analysis. Three objectives are listed in the decision framework for refined analysis: (1) model refinement; (2) identifying key sources of variability and uncertainty; and (3) identifying high exposure scenarios. For the latter objective, two classes of methods are introduced. Class *A* introduces methods that provide explicit measures for addressing high exposure scenarios. CART is the only method introduced in this class capable of addressing high exposure scenarios directly (Frey *et. al.,* 2003). Methods introduced in Class *B* (including ANOVA, conditional sensitivity analysis, and scatter plots) require judgment of an analyst for interpretation of the results and identification of inputs responsible for high exposure scenarios. The decision framework for method selection considering the first two objectives is illustrated in Figure 5-2.

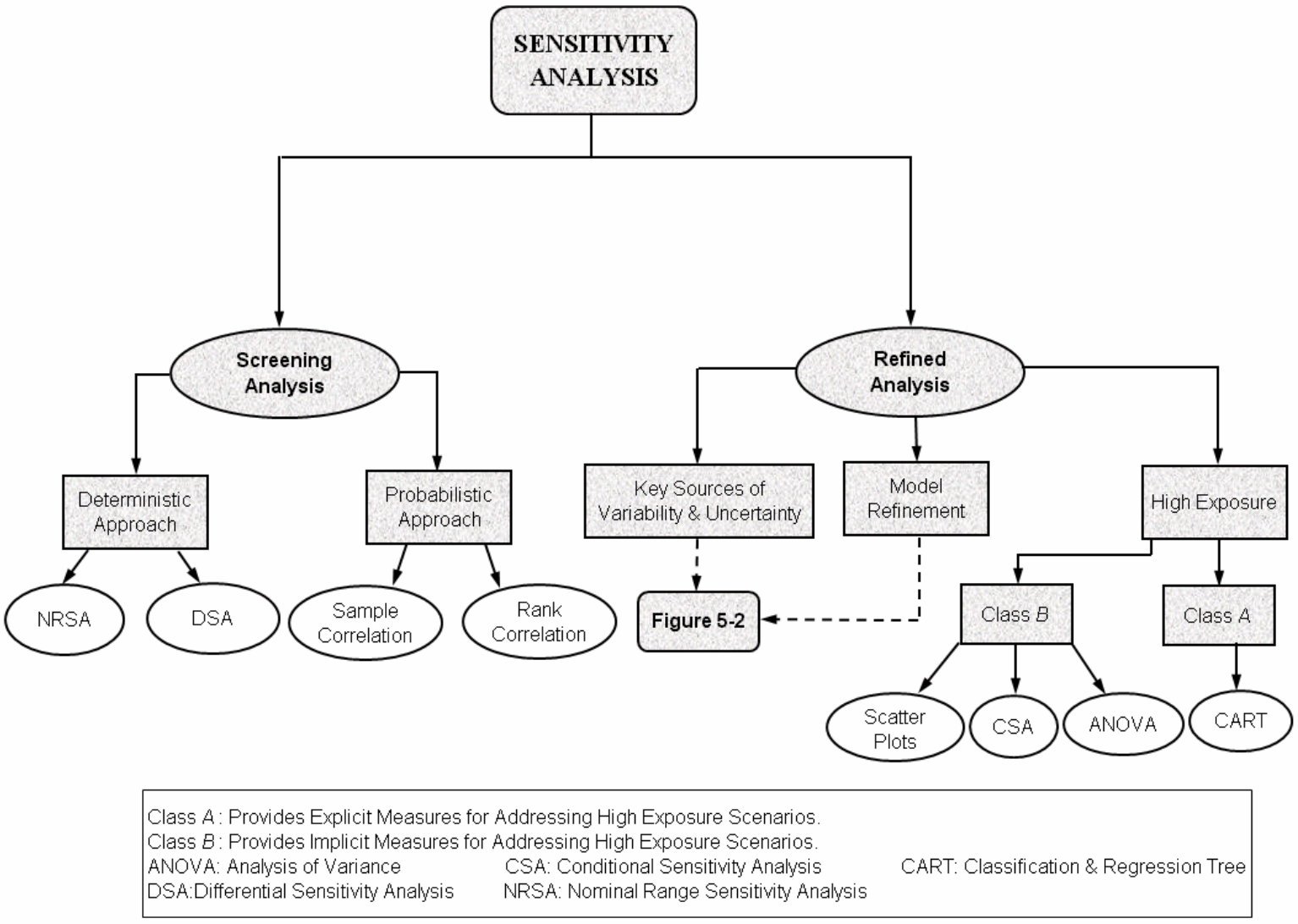

Figure 5-1. Decision Framework for Selecting an Appropriate Sensitivity Analysis Method.

When the objective of sensitivity analysis is either model refinement or identifying key sources of variability and uncertainty, the choice of an appropriate method for sensitivity analysis depends on the characteristics of the model under study. Figure 5-2 considers four characteristics for a food safety model: (1) non-linearity; (2) interaction; (3) categorical inputs; and (4) threshold and saturation points.

For non-linear models, the choice of a method first depends on whether or not the model is monotonic. The decision framework is further classified based upon the condition of whether a sensitivity analysis method is post hoc or integrated with software. Post hoc methods potentially can be integrated into the software framework, but need not be.

For models with interactions, the selection of method depends on whether the implementation of the selected method is post hoc. Post hoc methods are classified as of Class *C* or *D.* Methods in Class *C* directly take into account interaction between model inputs, while the results of methods in Class *D* involve analyst judgment in order to address interaction between inputs. Methods appropriate for models with categorical inputs are classified into two categories based upon post hoc or integrated implementation.

When the model under study has thresholds or saturation points, sensitivity analysis methods are classified based on whether they explicitly address such characteristic. Methods in class *C* including ANOVA and scatter plots (SP), which explicitly handle possible thresholds or saturation points in a model, while methods in Class *D* including CART and conditional sensitivity analysis, which require analyst judgment for quantifying available thresholds and saturation points.

With the help of the decision frameworks provided here, it is possible for an analyst to select an appropriate sensitivity analysis method for an application. For example, for a food safety model with many inputs, an analyst may decide to narrow the scope of sensitivity analysis by selecting a subset of inputs that control much of the output variation. Figure 5-1 presents recommended methods for screening level analysis. If an analyst prefers a deterministic approach to screening analysis and wants to account for the extreme values of inputs, NRSA is a good method to choose. After selecting a subset of inputs, an analyst may want to continue sensitivity analysis by identifying key sources of variability and uncertainty. For example, if the model is non-linear and monotonic, post hoc techniques could be used. Figure 5-2 recommends five methods for sensitivity analysis. If the model has non-linear and monotonic characteristics

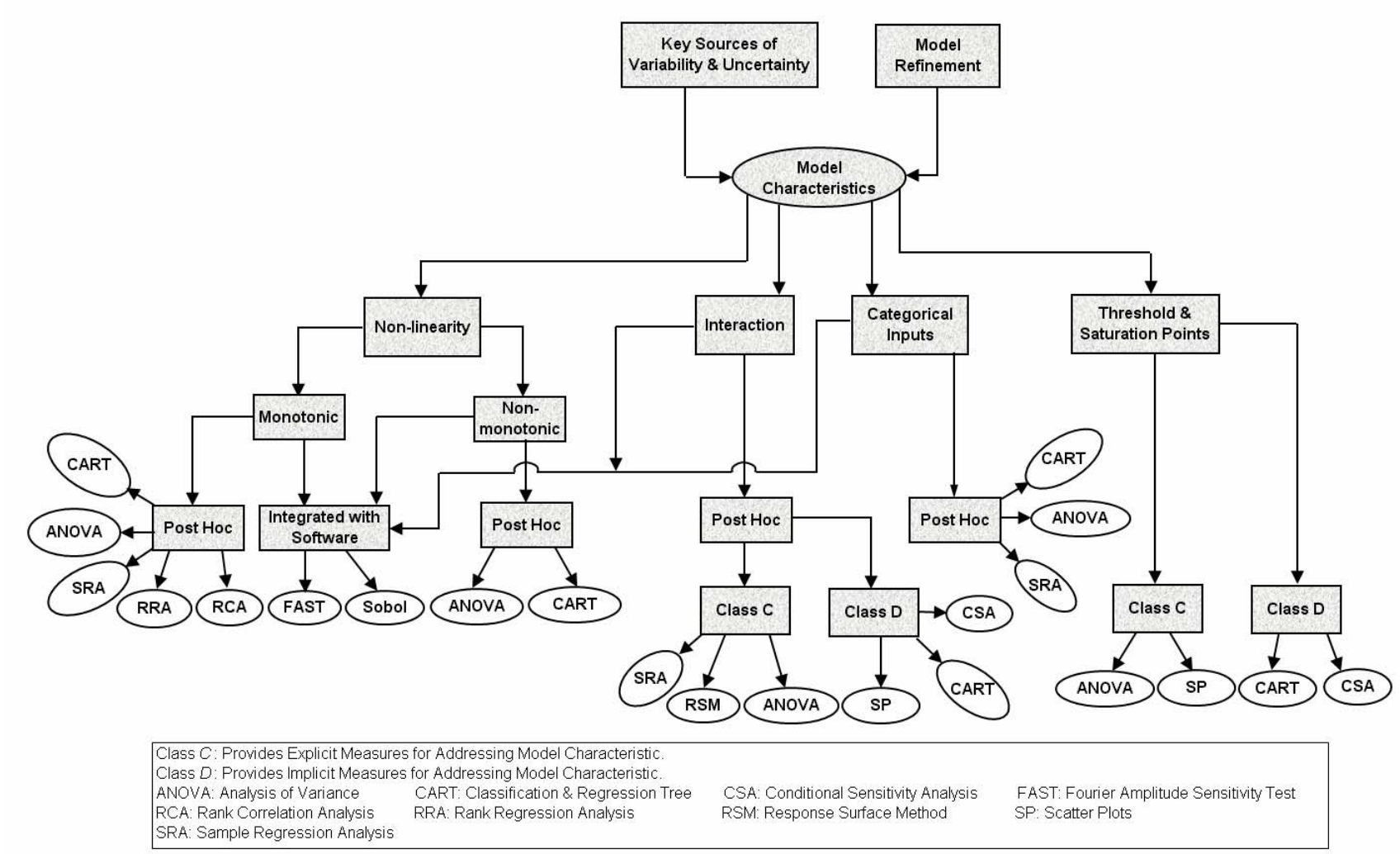

Figure 5-2. Decision Framework for Selecting Appropriate Sensitivity Analysis Method for Identifying Key Sources of Variability and Uncertainty and Model refinement as Key Objectives of the Analysis.

and contains categorical inputs and interactions between inputs, ANOVA and sample regression analysis are appropriate methods. An analyst may decide to apply both methods and compare the results to obtain more robust insights with respect to identification of key sources of variability and uncertainty in the model.

Table 5-1 summarizes key characteristics addressed by selected sensitivity analysis methods. These methods are evaluated by case studies performed by Frey *et al* (2003) and Patil and Frey (2003). Twelve characteristics listed in the table were evaluated based on those case study scenarios.

Ideally, a sensitivity analysis method should respond to the effects of simultaneous variation in all inputs. The methods were evaluated to determine if they address the nonlinearities in response to an input. The identification of the presence or absence of threshold in the model response was evaluated. The ability of sensitivity analysis methods to identify and provide detailed insights regarding the existence of interaction among inputs is considered as critical for food safety process risk models. Some methods handle only quantitative inputs and other can address both quantitative and categorical inputs.

Because high exposure cases are often of special interest, methods that can help identify and characterize conditions leading to high exposures may be preferable. The ability to address uncertainty about importance of inputs via two-dimensional analysis provides robustness and confidence in the control measure applied based upon insights from the analysis. Some methods are easier to apply in practice than others. The ease of application may often constrain the feasibility of a method. A method is typically easier to implement when software tools for implementing the method already exist, especially if they have user-friendly interfaces. For example, Pearson correlation coefficients are easy to implement for users of Crystal Ball™. Of course, ease of implementation will be a function of software availability and programming skill level.

The ability to produce quantitative rankings and the ability to evaluate the statistical significance of the rankings are useful to identify the relative importance of inputs and the confidence that should be imputed to the rankings. Some methods produce more useful measures by which to discriminate the importance among similarly ranked inputs. Finally, although each method has a different theoretical basis, the bottom line for practical use of the methods is

Table 5-1. Summary of Key Characteristics of Selected Sensitivity Analysis Methods

Characteristic	Sensitivity Analysis Method													
	NRSA	DSA	Correlation		Regression		ANOVA	CART	FAST	Sobol	RSM	MII	SP	CSA
			Sample	Rank	Linear	Rank								
Simultaneous Variation	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Non-linearity	No	No	No	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Threshold	No	No	No	No	No	No	Yes [c]	Yes	No	No	No	No	Yes	Yes
Interaction	No	No	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Qualitative vs. Quantitative inputs	No	No	No	No	Yes	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
High Exposure	No	No	No	No	No	No	Yes	Yes	No	No	No	No	Yes	Yes
Two-Dimensional Analysis	No	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	No	No
Ease of Implementation	No	Yes	Yes	Yes	Yes	Yes	No	No	No	No	No	No	Yes	No
Quantitative Ranking of Inputs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No[a]	Yes	Yes	Yes	Yes	No	No
Measure of Statistical Significance	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No
Discrimination of Important Inputs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No [b]	No
Robust in Practice	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

[a] A method for ranking the input based upon the contribution of each input to reduction in total deviance was explored and is promising for future development

[b] Can be based upon expert judgment

[c] Depends on proper definition of factor levels

NRSA: Nominal Sensitivity Analysis DSA: Differential Sensitivity Analysis ANOVA: Analysis of Variance CART: Classification & Regression Tree
CSA: Conditional Sensitivity Analysis RSM: Response Surface Method MII: Mutual Information Index SP: Scatter Plots
FAST: Fourier Amplitude Sensitivity Test

whether they produce reasonable results even if there are departures from key assumptions of the method.

5.3 Available Methods for Sensitivity Analysis

This section presents a brief discussion regarding available methods for sensitivity analysis which can be applied to food safety risk assessment models. Frey *et al.* (2003) and Patil and Frey (2003) demonstrated the application of selected sensitivity analysis methods to three food safety process risk models. These methods are briefly discussed in the following, and key advantages and disadvantages for each method are highlighted. These methods include nominal range sensitivity analysis (NRSA), differential sensitivity analysis (DSA), standardized regression analysis, rank regression, analysis of variance (ANOVA), sample and rank correlation coefficients, classification and regression tree (CART), scatter plots, and conditional sensitivity analysis (CSA). These techniques are classified into three categories, which are mathematical, statistical, and graphical sensitivity analysis methods. The three categories are explained in Sections 5.3.1 through 5.3.3, respectively. Four other statistical sensitivity analysis methods that are not commonly applied to food safety process risk models are discussed in Appendix *A*. These methods include FAST, response surface method (RSM), MII, and Sobol method. Appendix *A* provides a discussion regarding description, application, and interpretation of results for each of the four methods.

An additional category of sensitivity analysis methods that may be of interest to practitioners is Bayesian methods. Such techniques are complex and beyond the scope of this document. The interested reader is referred to Saltelli *et al* (2000) for more information.

5.3.1 Mathematical Methods for Sensitivity Analysis

Mathematical methods typically address the local or linear sensitivity of the output to perturbations or ranges of individually varied inputs. These methods do not address the variance in the output due to the variance in the inputs (Morgan and Henrion, 1990). These methods are helpful in screening the most important inputs (e.g., Brun *et al.*, 1997). Among several mathematical methods discussed by Frey and Patil (2002), two mathematical methods were selected and applied to the case studies defined for selected food safety process risk models (Frey *et al.*, 2003). The two methods are nominal range sensitivity analysis (NRSA) and differential sensitivity analysis (DSA). The two methods are briefly discussed in Sections 5.3.1.1 and 5.3.1.2.

5.3.1.1 Nominal Range Sensitivity Analysis Method

NRSA is known as local sensitivity analysis or threshold analysis (Cullen and Frey 1999; Critchfield and Willard, 1986). This method is applicable to deterministic models. A typical use of NRSA is as a screening analysis to identify the most important inputs in a model in a probabilistic framework (Cullen and Frey, 1999).

NRSA is used to evaluate the effect on a model output of varying only one of the model inputs across its entire range of plausible values, while holding all other inputs at their nominal or base-case values (Cullen and Frey, 1999). The difference in the model output due to the change in the input is referred to as the sensitivity to the particular input (Morgan and Henrion, 1990). The sensitivity analysis can be repeated for any number of individual model inputs.

The results of NRSA are most valid when applied to a linear model. In such cases, it would be possible to rank the relative importance of each input based upon the magnitude of the calculated sensitivity indices as long as the ranges assigned to each input are accurate. For models with interaction terms, the sensitivity of the output to a given input may depend on interactions with other inputs. NRSA cannot characterize such interactions. Thus, for models with interactions, NRSA may not provide a reliable rank ordering of key inputs.

5.3.1.2 Differential Sensitivity Analysis Method

DSA is a local sensitivity analysis method. It calculates the sensitivity of the output to small deviations in the point estimate of an input (Frey *et al.,* 2003). DSA is performed with respect to some point in the domain of the model. A small perturbation with respect to the point value of each model input on an individual basis, such as a change of plus or minus one percent, can be used to evaluate the corresponding change in the model output.

DSA is conceptually easy to apply. The computational time is proportional to the number of inputs. DSA is especially useful when a high degree of confidence is attributed to a point estimate and thus the variation in the output need only be tested for small variations around the point estimate.

DSA presumes a linear model response and deals only with small perturbations. DSA may provide unreliable insights if the range of variation among inputs varies disproportionately to their local sensitivity or if the model response is not linear over the domain of interest.

5.3.2 Statistical Methods for Sensitivity Analysis Methods

An advantage of statistical methods is that they can identify the effect of simultaneous interactions among multiple inputs. Distributions for model inputs can be propagated through a model using a variety of sampling techniques, such as Monte Carlo simulation, Latin Hypercube sampling, and other methods (Cullen and Frey, 1999).

Among several statistical methods discussed by Frey and Patil (2002), five methods were selected and applied to two food safety process risk models (Frey *et al.,* 2003). FAST, RSM, and MII were evaluated by Patil and Frey (2003). The methods summarized here in Sections 5.3.2.1 to 5.3.2.4 include: (1) sample and rank correlation analysis; (2) sample and rank linear regression; (3) ANOVA; and (4) CART, respectively. As a caveat, the terminology used here for sample correlation analysis and sample regression analysis methods are not necessarily universally accepted usage. These terms are used as a convenient way to distinguish analyses based upon the sample values from analyses based upon the ranks of sample values.

5.3.2.1 Sample and Rank Correlation Analysis

The correlation coefficient method is typically classified into two types: (1) sample or Pearson; and (2) rank or Spearman. The sample (Pearson) correlation analysis evaluates the strength of linear association between paired input and output values. The rank (Spearman) correlation analysis is based upon ranks of sample data, and is a measure of the strength of the monotonic relationship between two random variables. Thus, it can account for monotonic nonlinear relationships (Gibbons 1985, Siegel and Castellan 1988, and Kendall 1990).

Both sample and rank correlation coefficients can range from -1 to +1. A value of zero represents a lack of correlation (Edwards, 1976). The statistical significance of sample and rank correlation coefficients can be evaluated based upon the use of an inverse Fisher transformation (Steel et. al., 1997).

A disadvantage of correlation coefficient methods is that the existence of correlation does not imply causation. In addition, sample coefficients are inaccurate for nonlinear models, and rank coefficients are inaccurate for non-monotonic models. Neither sample nor rank coefficients can capture complex dependencies such as thresholds or directly deal with interactions among inputs.

5.3.2.2 Sample and Rank Regression Analysis

Regression analysis can be employed as a probabilistic sensitivity analysis technique as demonstrated by Iman et al. (1985). Two types of regression analysis are considered here: (1) sample regression analysis; and (2) rank regression analysis. Sample regression involves fitting a model to a dataset that includes input values and corresponding output values. Typically, the data for each input and output are standardized to remove the effects of scale. Rank regression is based upon the ranks for the inputs and the output (Neter et. al., 1996). Rank regression is especially useful when there is high amount of variance or noise in the data or if the model is non-linear but monotonic.

For both sample and rank regression analysis, the effect of variation of inputs on the variation in output can be evaluated using regression coefficients, standard errors of regression coefficients, and the level of significance of the regression coefficients (Devore and Peck, 1996; Steel et al., 1997; Sen and Srivastava, 1990). For a linear model, a standardized regression coefficient provides a measure of the relative importance of each input (Devore and Peck, 1996; Neter et al., 1996; Iman et al., 1985). Similarly, rank regression coefficients, which are independent of scale, can be used to rank the inputs. Rank regression coefficients cannot be transformed back to obtain sensitivities in terms of the original ranges of an input.

The 95 percentile confidence intervals for both sample and rank regression coefficients can be used to evaluate the degree of ambiguity of estimated ranks (Frey *et al.,* 2003). The adequacy of the regression model can be assessed using the coefficient of multiple determination, R^2, which is a measure of the amount of variance in the dependent variable explained by the model in the regression analysis (Draper and Smith, 1981). In rank regression analysis, a high R^2 value indicates a monotonic relationship.

The key potential drawbacks of regression analysis include: (1) possible lack of robustness if key assumptions of regression are not met; (2) the need to assume a functional form for the relationship between an output and a selected input; and (3) potential ambiguities in interpretation. Because regression analysis is not model independent, a proper characterization of interactions and non-linearities requires *a priori* specification of an appropriate model.

5.3.2.3 Analysis of Variance

ANOVA can be applied to models that are linear, non-linear, monotonic, or non-monotonic. ANOVA can address both qualitative and quantitative inputs (Steel et. al., 1997).

An input is referred to as a "factor" and specific ranges of values for each factor are considered as factor "levels" in ANOVA. In ANOVA a "treatment" is a specific combination of levels for different factors. An output is referred to as a "response variable". A "contrast" is a linear combination of two or more factor level means. For example, a contrast can be specified to evaluate the mean growth of pathogen organisms when the storage temperature varies between high and low levels for a specific storage time.

ANOVA tests the hypothesis that the means among two or more input levels are equal, under the assumption that the output for each of the input levels is normally distributed with the same variance (Neter et. al., 1996). ANOVA uses the F test to determine whether there exists a significant difference among treatment means or interactions between factors. The F values can be used to rank the factors based on their relative magnitude (Carlucci, 1999). The higher the F value for a factor is, the more sensitive the output is to the factor. Therefore, factors with higher F values are given higher rankings.

In a probabilistic setting, the F value is estimated based upon a random sample of values for factors. Hence, the F value may be treated as a random variable. In order to discern whether the relative importance of different factors can be discriminated, uncertainty in estimated F values should be quantified. The method of bootstrap simulation can be used to generate sampling distributions of uncertainty for F values. Bootstrap simulation is a numerical method for estimating confidence intervals of statistics (Efron and Tibshirani, 1993). Quantifying the uncertainty in F values identifies the magnitude of difference in F values of two factors that indicates a clear difference in rankings (Frey *et. al.,* 2003).

5.3.2.4 Classification and Regression Tree (CART)

CART is a method for partitioning the output values based on classification rules that depend on cut-off values of selected inputs. CART produces partitioned datasets with statistically significantly different mean values for an output. Hence, CART can provide insight into conditions that lead to high exposure or risk. CART is applicable to linear or non-linear models, including models with interactions and thresholds.

CART uses dendritic terminology. In tree-based models, there are branches, branch splits or internal nodes, and leaf or terminal nodes (Washington et al., 1997). The components of a regression tree are shown in Figure 5-3. A node refers to a cut-off value. Nodes can be a root

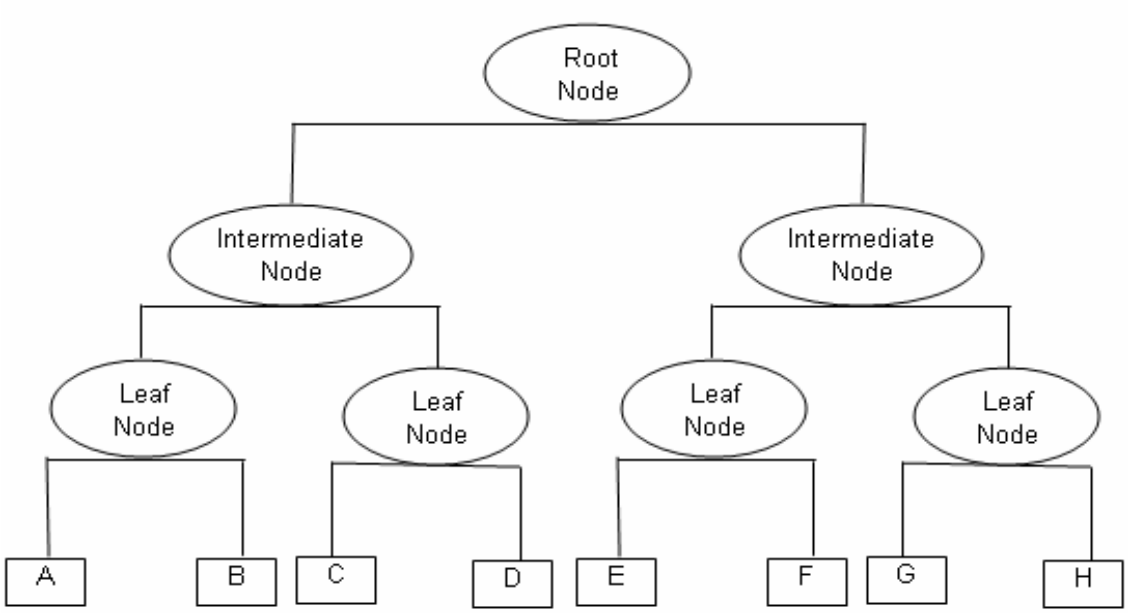

Figure 5-3. Schematic Diagram of a Classification and Regression Tree Illustrating Rout Node, Intermediate Nodes, and Terminal Leaves.

node, intermediate nodes, or leaf nodes. A root node is the node at which the data is first split. Intermediate nodes are the nodes at which the data is successively split. Leaf nodes represent the penultimate data split. Branches are the conditions on the input variables that determine the classification rules. A set of conditions on the input variable from the root node leading to a root node is called a path or classification rule.

Inputs selected in a regression tree have significant effects on the response variability. As an indication of priority among selected inputs in a regression tree, inputs can be prioritized based upon their position in the tree. Typically those inputs in the top nodes are more important than inputs in the lower nodes. Another indication of input importance is if it is selected repeatedly at multiple levels within the tree.

An alternative sensitivity index is the total contribution of selected inputs to reduction in the total dataset deviance can be used. Hence, the higher the contribution of the input to the reduction of the total deviance, the higher is the rank (Frey *et al.,* 2003).

There are several advantages to CART. CART is a nonparametric method that does not require assumptions of a particular distribution for the error term. CART is more resistant than other methods to the effects of outliers since splits usually occur at non-outlier values (Roberts et al., 1999). Unless there is a collinearity problem, a regression tree selects only the most

important inputs and values of these inputs that result in the maximum reduction in deviance (Hallmark et al., 2002). Moreover, application of discrete, continuous, and also qualitative inputs is possible in CART.

There are some disadvantages to CART. Because CART is not a standard analysis technique, it is not included in many major statistical software packages (Levis, 2000). Moreover, there are alternative ways to prioritize inputs based on the results of the CART, which requires judgment on the part of the analyst.

5.3.3 Graphical Methods for Sensitivity Analysis

Graphical methods can be used as a screening method to identify the existence of complex dependencies between inputs and outputs in order to help select appropriate statistical sensitivity analysis methods (e.g., McCamly and Rudel, 1995). Graphical methods can be used as a complement to mathematical and statistical methods to better interpret sensitivity analysis results (e.g., Stiber et al., 1999; Critchfield and Willard, 1986).

Two graphical methods -scatter plots and conditional sensitivity analysis- are briefly discussed in Sections 5.1.3.1 and 5.1.3.2, respectively.

5.3.3.1 Scatter Plots

Scatter plots are used for visual assessment of the influence of individual inputs on an output (Cook, 1994; Galvao et al., 2001). A scatter plot is often used to explore patterns between inputs and outputs of a probabilistic simulation. Scatter plots are also often used as a first step in other analyses, such as regression analysis and response surface methods. The scatter plot displays paired sample values for an input and output and can enable an analyst to gain insight into the general trend between them.

A key advantage of scatter plots is that they may reveal potentially complex dependencies, such as non-linearity, interactions, or thresholds. A potential disadvantage of scatter plots is that they may be tedious to generate if there are a large number of inputs and outputs, unless an applicable software package is used to automatically generate multiple scatter plots. In addition, the interpretation of scatter plots is often qualitative, and may rely on judgment of an analyst.

5.3.3.2 Conditional Sensitivity Analysis Method

Conditional sensitivity analysis is considered to be a graphical method since the results from the method are often presented in form of graphs. Conditional sensitivity analysis involves

evaluating the effect of changes in a subset of model inputs while other inputs are held at fixed values. The motivation for using this technique is that the effect of variation in any input on the output in a non-linear model cannot be adequately captured by mathematical methods such as NRSA.

Conditional sensitivity analysis can deal with multiple combinations of point values among the inputs. The response is calculated for point values of the selected input variable at arbitrarily spaced intervals or for randomly generated points. The purpose of conditional sensitivity analysis is to cover the full scope of the variation of the selected variables. A graph is plotted from these data points showing the response curve for a specific variable conditional on the fixed values of remaining variables (e.g., minimum, mean, and maximum values). The process is repeated for other values of the other input variables.

Non-linearity, saturation points and thresholds can be identified based upon conditional sensitivity analysis. These insights are under assumptions that other variables are fixed at particular values. The drawback for the method is that the response curve may provide insufficient information to rank inputs. For example, if two inputs have non-linear response, it may be difficult to tell which one has a higher degree of variance.

5.4 Summary

The selection of sensitivity analysis methods depends on factors such as objective of the analysis, characteristics of the model under study, amount of detail expected from sensitivity analysis, characteristics of the software used for sensitivity analysis, and available computing resources. This chapter provided a series of key questions and brief discussions regarding the insight that an analyst may gain by addressing those questions. A practitioner should provide answers to those questions before choosing a method. Based on the key questions and discussions provided in this chapter, decision frameworks summarizing the discussions regarding selection of appropriate sensitivity analysis methods were introduced. Finally, this chapter gave a brief review of available sensitivity analysis methods that can be used in the food safety process risk models. Additional methods are summarized in Appendix A.

The next chapter introduces the procedures for implementing the selected sensitivity analysis methods.

6 PROCEDURE FOR APPLICATION OF SENSITIVITY ANALYSIS METHODS

The objective of this chapter is to provide procedures for application of selected sensitivity analysis methods to food safety risk assessment models. An assumption here is that the dataset required for sensitivity analysis has been generated or that the sensitivity analysis method is available as an integral part of the model. Hence, the application of the selected sensitivity analysis method can be either post hoc or an integral part of the analysis, as appropriate to the method.

For each sensitivity analysis method, the procedure for application is provided, typically in the form of a flow diagram. Furthermore, a discussion is presented for each method highlighting key considerations regarding application of the method.

This chapter describes procedures for application of the following sensitivity analysis methods: (1) NRSA; (2) DSA; (3) sample and rank correlation; (4) sample and rank regression; (5) ANOVA; (6) CART; (7) scatter plots; and (8) conditional sensitivity analysis. Appendix *A* provides a brief description of other sensitivity analysis methods that are less commonly used but that are potentially useful for application to food safety risk assessment models. These methods include FAST, Sobol's method, MII, and RSM.

The procedures for most of the methods included here are based upon case studies by Frey *et al.* (2003) for the *E. coli* and *Listeria* food safety risk assessment models. As described in previous chapters, the methods are grouped into three categories: mathematical, statistical, and graphical methods. The procedures for application of these methods are presented in Sections 6.1, 6.2, and 6.3, respectively. Section 6.4 provides a summary for the chapter.

The procedures introduced in this chapter for implementing sensitivity analysis methods are independent of any specific statistical software package. The only exception is that the procedure presented for CART in Section 6.2.4 is based on the S-PLUS^{TM} software package.

6.1 Procedure for Application of Mathematical Sensitivity Analysis Methods

In this section, the procedures for application of DSA and NRSA are presented. The procedures are similar for both methods. These two methods are described in Section 5.3.1 and have been evaluated by Frey *et al* (2003).

Figure 6-1 is a schematic diagram of the procedure for application of NRSA and DSA. Because DSA and NRSA are local sensitivity analysis methods, the first step is to identify the

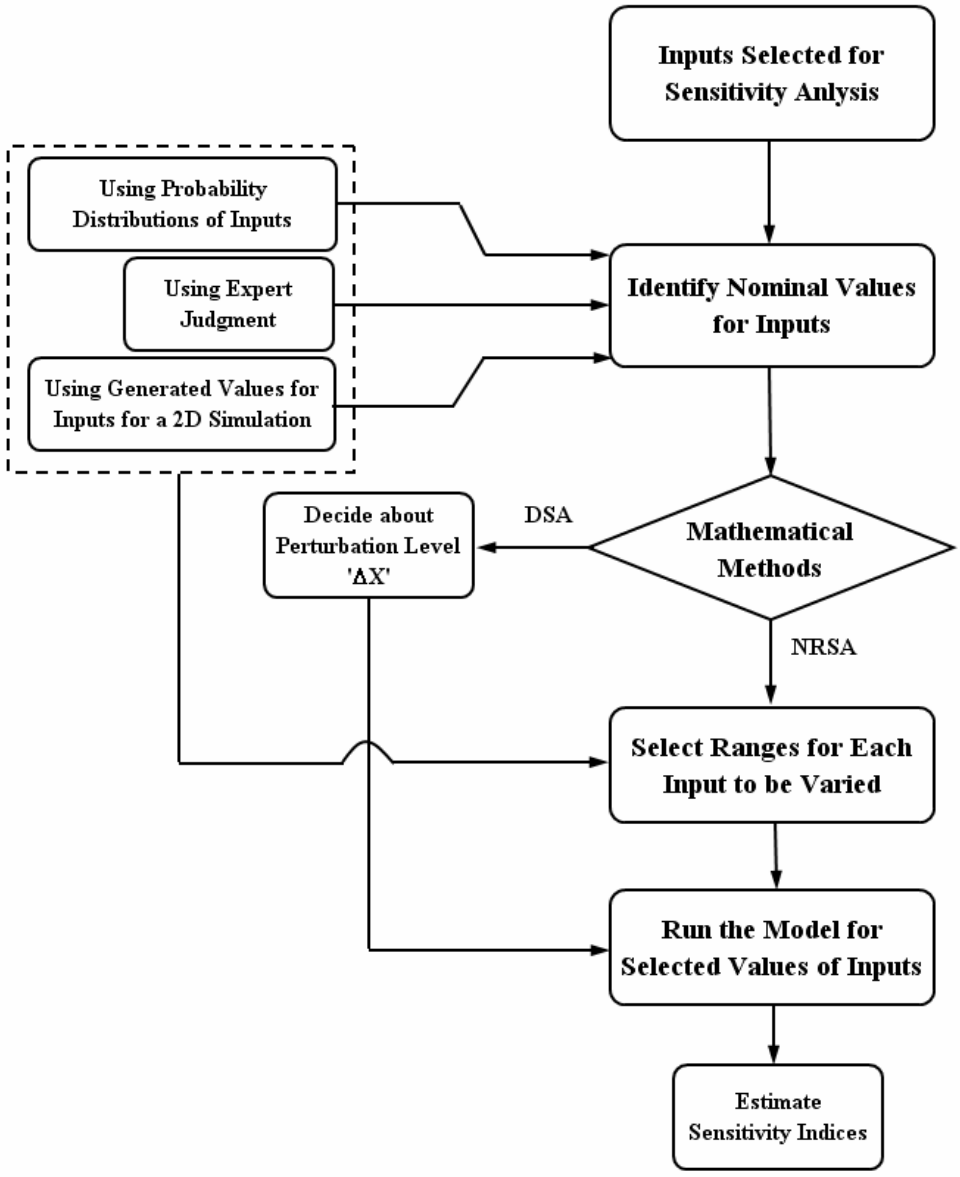

Figure 6-1. Schematic Diagram for Procedure of Application of NRSA and DSA.

point values for all model inputs that will be the basis of the analysis. The selection of a nominal point is dependent upon the assessment objectives. A nominal point could be intended to represent the mean, median, minimum, maximum, or a particular percentile of a distribution for a given input. NRSA differs from DSA primarily in that a wider range of variation is evaluated for each input when performing the sensitivity analysis. Thus, the nominal point chosen for NRSA is typically a central value such as a median or mean. However, this is not a requirement of the method, but a convention.

In DSA and NRSA, the ranges for the inputs are treated as perturbations or intervals rather than distributions. Thus, if these methods are applied to a model for which probability distributions have been assigned to inputs, the choice of a nominal point can be based upon the distributions. For NRSA, the choice of ranges of variation for each input might typically be based upon a 95 percent probability range. Other ranges, such as the difference between the maximum and minimum values may be used. However, this is not a requirement but a matter of preference by the analyst. For those inputs that incorporate both variability and uncertainty in a two-dimensional probabilistic framework, the selection of minimum and maximum values could be based upon collapsing all of the randomly simulated values in the two-dimensions into a single dimension, and then choosing endpoints based upon the desired (based upon the analyst's preference) proportion of simulated values that should be enclosed by the interval.

For DSA, once the nominal values are obtained for each input, a choice should be made regarding the perturbations that should be made to each input. A strict application of DSA would involve calculating the first partial derivative for each input. A practical alternative is to specify some small percentage variation for each input, such as one percent change, as the basis for estimating the corresponding change in the model output. An analyst could vary the percentage (e.g., for 1% to 5%) to test the accuracy of the approximation and sensitivity of the results to the perturbation level.

For NRSA, each input is varied one at a time over its specified range. The range of variation of the inputs for NRSA is typically larger than that for DSA. For both methods, the corresponding range of variation in the output is estimated. Typically, the range of variation in the output is divided by the nominal value of the output to arrive at a normalized measure of sensitivity with regard to each input.

6.2 Procedure for Application of Statistical Sensitivity Analysis Methods

In this section, the procedure for application of statistical sensitivity analysis methods is presented. These methods include: (1) regression analysis; (2) correlation analysis; (3) ANOVA; (4) CART. The procedures for application of these methods are described in Sections 6.2.1 to 6.2.4, respectively. Each of these methods has been briefly described in Section 5.3.2.

6.2.1 Regression Analysis

In this section the procedure for application of regression analysis is introduced. Two types of regression analysis are considered here: (1) sample regression analysis; and (2) rank

regression analysis. Because the procedures for the two methods are similar, a single flow diagram shown in Figure 6-2 is provided. This figure provides procedures for a practitioner to apply either sample or rank regression analysis methods.

The key assumption for application of regression analysis is that a dataset, including values from model inputs and corresponding output values is available. For the case of rank regression, both output and input values are rank ordered before the regression analysis is implemented. For a one-dimensional analysis incorporating only variability or uncertainty, or for a scenario that variability and uncertainty are co-mingled into a single dimension, a single dataset should be constructed for regression analysis. In contrast, when an analyst wants to fully distinguish between variability and uncertainty multiple datasets should be constructed in a two-dimensional probabilistic approach. Specifically, the multiple datasets could represent variability for different uncertainty realizations or uncertainty for different variability iterations. Alternative probabilistic approaches (i.e., one-dimensional versus two-dimensional) are discussed in Section 4.4.

In Figure 6-2, *m* refers to the number of datasets constructed in a probabilistic simulation. For a one-dimensional analysis *m* is equal to one, while *m* is the number of uncertainty realizations or variability iterations for a two-dimensional analysis.

As shown in Figure 6-2, the first step is to choose between sample and rank regression analysis. If rank regression is selected, each dataset should first be rank-ordered. Some statistical software packages can rank the dataset automatically (e.g., SAS©); otherwise, the analyst may have to assign ranks by performing the calculations with the assistance of a spreadsheet or other software. If sample regression analysis is selected, the dataset should be normalized. A dataset is typically normalized based upon the mean and standard deviation of the data. There are alternative normalization techniques, and the one used by Frey *et al.* (2003) is documented by Neter *et al.* (1996).

The next step is to select a functional form for the regression model. A practitioner may choose a linear model as a starting point and change the functional form of the regression model later following the procedure provided below. There are also standardized procedures such as forward, backward, and stepwise selection of input variables in a regression model that can be used (Neter *et al.,* 1996). The regression model also may include interaction terms and higher order terms

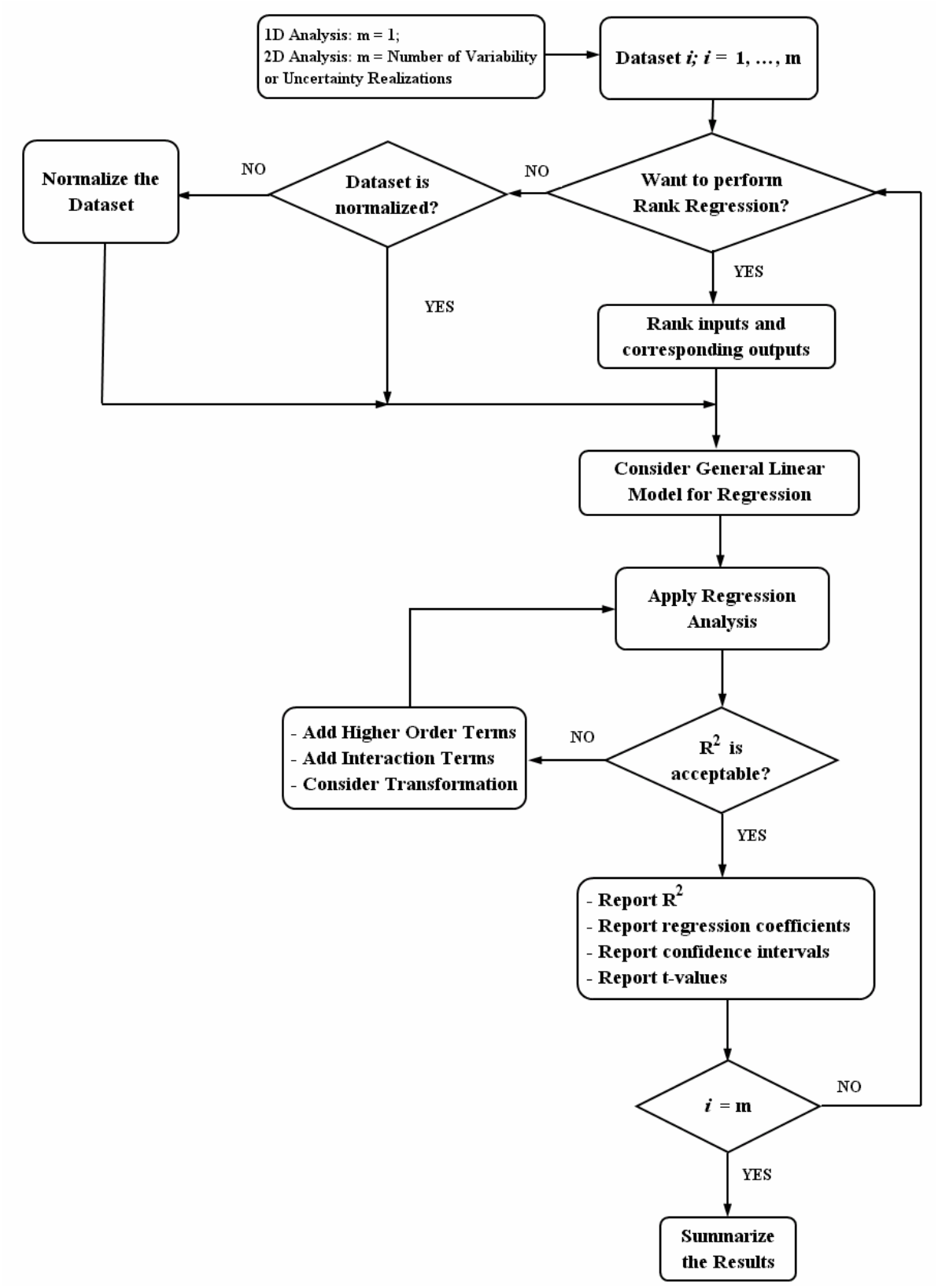

Figure 6-2. Schematic Diagram for Procedure of Application of Regression Analysis.

initially (e.g., quadratic, and cubic terms) based upon judgment and characteristics of the model that are already known.

After determining the functional form of a regression model, the next step is to apply regression analysis to estimate the coefficient of determination, R^2. The R^2 value indicates the percent of variation in the model response that is explained by the inputs considered in a regression model. R^2 can be used as a diagnostic to evaluate the goodness-of-fit for a fitted regression model. A high R^2 value implies that the assumption for the functional relation between the output and inputs is substantially valid. In contrast, low R^2 values indicate that the underlying assumption may not be valid and the fitted model is not good enough. Interactions and higher order terms, such as quadratic, and cubic terms may be necessary to improve the goodness-of-fit. An alternative approach for improving the goodness-of-fit is the application of a linearizing transformation of the dataset (Neter *et al.,* 1996). The drawback of the transformation technique is that the results of the analysis will be based on the transformed dataset and not the original untransformed dataset.

For one-dimensional cases, it is practical to modify the regression model form until a satisfactory value for R^2 is obtained. However, for two-dimensional probabilistic frameworks, the same functional form of the regression model should be applied since it is not practical to develop separate regression models for each realization of uncertainty or variability. Alternatively, an analyst can apply the functional form selection process to one of the *m* datasets selected in random, and evaluate the distribution of R^2 obtained from applying regression analysis with the selected functional form to the other *m-1*datasets.

A regression coefficient is estimated for each input considered in the regression model. A test statistic can be calculated to evaluate whether each estimated coefficient is statistically significant. Details of calculation of the regression coefficients and test statistics are given in Frey et al. (2003). Estimated regression coefficients for standardized inputs can be reported as sensitivity indices for rank-ordering the importance of the inputs. Specifically, the inputs can be prioritized based upon the relative magnitude of statistically significant coefficients. As an alternative approach, the amount of explained variability corresponding to an input effect, including simple and interaction effects can be used as a sensitivity index (Rose *et al.,* 1991).

6.2.2 Correlation Analysis

This section introduces a procedure for application of correlation coefficients for sensitivity analysis. Two types of correlation analysis are considered here: (1) sample (Pearson) correlation coefficients; and (2) rank (Spearman) correlation coefficients. Since the procedures for both methods are similar, a single flowchart presenting the procedures is provided in Figure 6-3. A key assumption for application of correlation analysis is that a dataset including paired output and input values is available. For rank correlation analysis, the inputs and outputs are rank-ordered before applying the method. Some statistical software packages can rank the dataset automatically (e.g., SAS©); otherwise, the analyst may have to assign ranks by performing the calculations with the assistance of a spreadsheet or other software packages.

The correlation analysis may be applied to a one-dimensional or two-dimensional probabilistic framework. Thus, a practitioner may have a single dataset obtained from a one-dimensional simulation or multiple datasets representing either variability for multiple realizations of uncertainty, or uncertainty for multiple realizations of variability. For each dataset, the correlation coefficients and corresponding P value can be calculated. The calculated statistic can be used to test the hypothesis that the true (population) correlation coefficient is different from zero, indicating whether the coefficient is statistically significant. For one-dimensional frameworks, a single correlation coefficient and corresponding P value would be calculated for each input. For two-dimensional frameworks multiple correlation coefficients and corresponding statistics would be calculated for each input for each realization of the model.

6.2.3 Analysis of Variance

This section introduces a procedure for application of ANOVA for sensitivity analysis. Figure 6-4 presents a procedure for using ANOVA as a sensitivity analysis method.

As shown in Figure 6-4, the first step for application of ANOVA is to identify whether any of the factors are continuous and to determine the levels for factors if continuous factors exist. If an input is categorical, the nominal values are treated as factor levels. For continuous inputs, generated values for each input in a probabilistic simulation can be partitioned into defined intervals to create factor levels (Kleijnen *et al.*, 1999). Frey *et al.* (2003) demonstrated three approaches for defining factor levels for continuous inputs: (1) equal intervals; (2) equal percentiles; and (3) visual inspection of the cumulative distribution function (CDF) for each

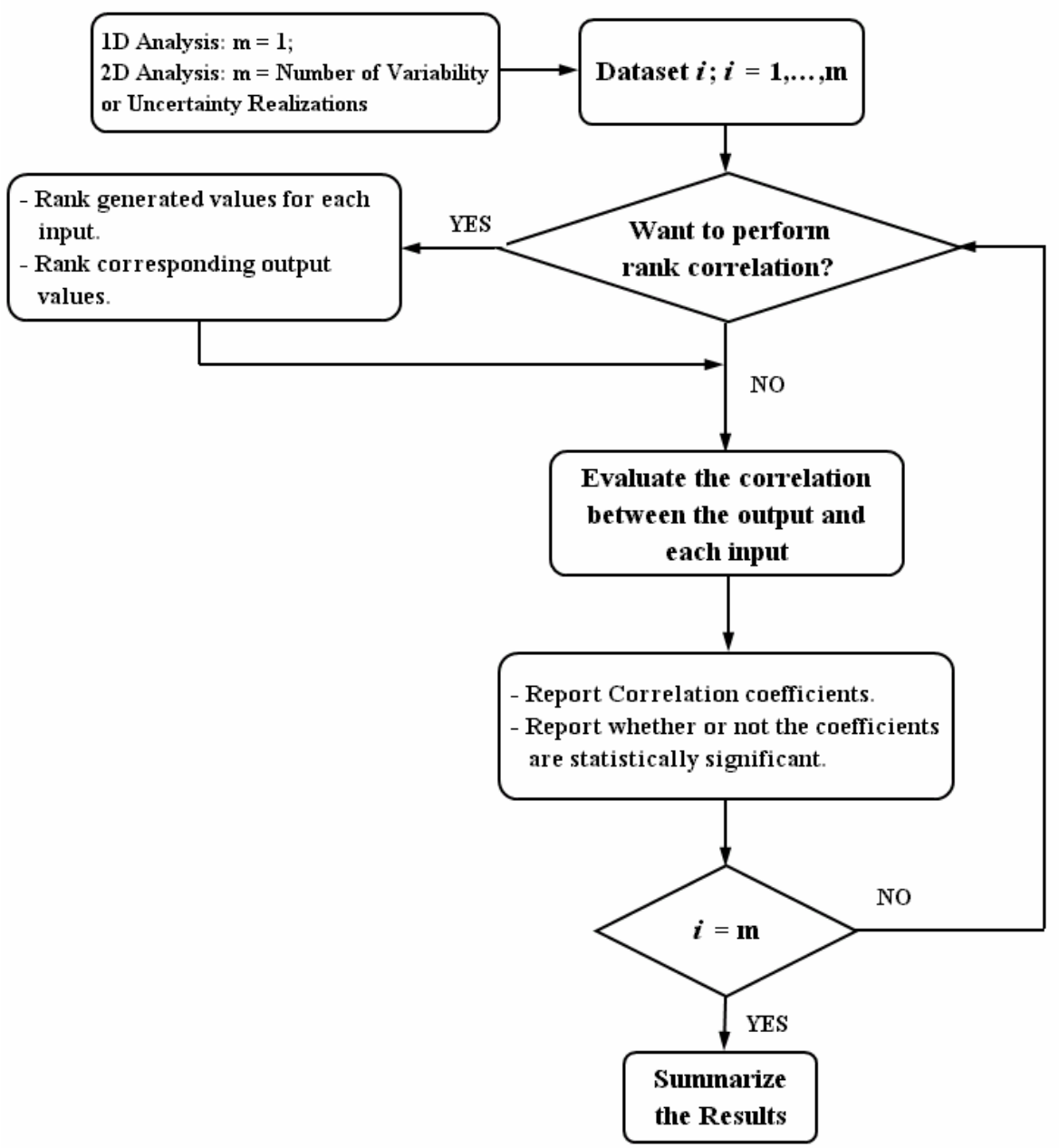

Figure 6-3. Schematic Diagram for Procedure of Application of Sample and Rank Correlation Coefficient Methods.

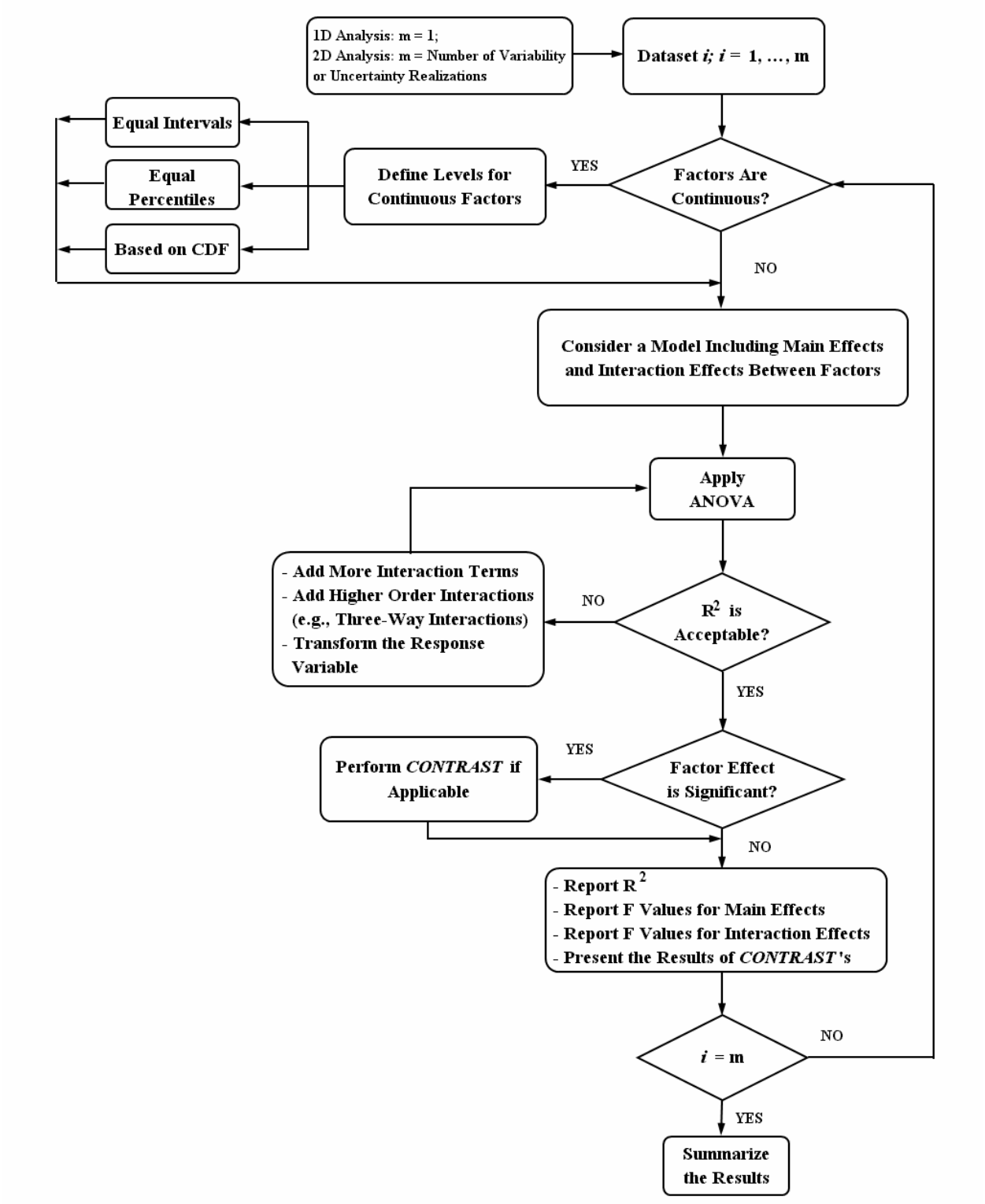

Figure 6-4. Schematic Diagram for Procedure of Application of ANOVA.

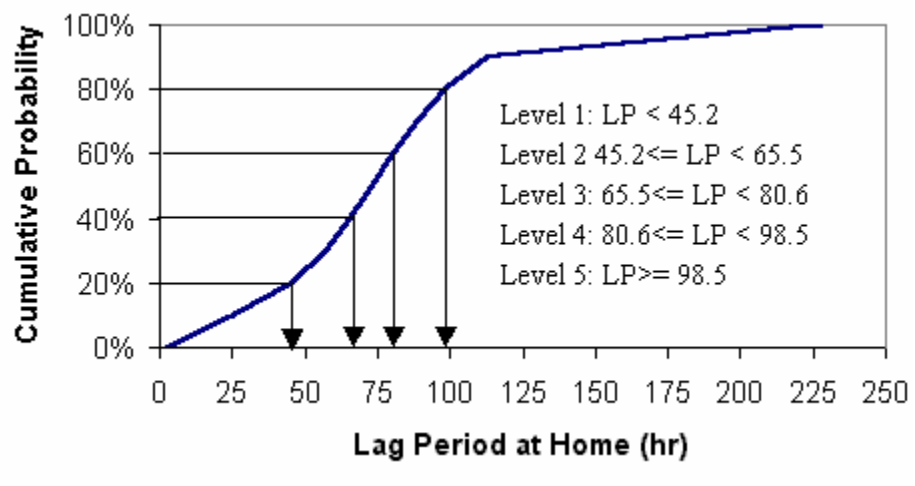

Figure 6-5. Definition of Levels for Lag Period Based Upon Equal Percentiles.

input.

For the equal intervals approach, each input domain is classified into equal ranges. For the equal percentiles, the CDF of the generated values for an input can be used for defining factor levels at equal percentiles. A practitioner can use visual inspection of the CDF for generated values of an input in order to define boundaries for each factor level corresponding to percentiles of the CDF that indicate a substantial change in the shape (e.g., inflection point). These methods are further explained with some examples.

Figure 6-5 shows an example in which factor levels are defined for lag period at home, which is a continuous input, using equal percentiles. In this figure, five intervals are arbitrarily defined using equal percentiles corresponding to 20^{th}, 40^{th}, 60^{th}, and 80^{th} percentiles of the CDF. The original number of samples generated from probability distribution of lag period was 10,000. Hence, approximately 2,000 samples lie within each factor level. The number of samples simulated affects the statistical significance tests. Criteria for minimum required number of samples in each level are discussed by Giesbrecht and Gumpertz (1996). The number of factor levels also impacts the results of ANOVA and computational effort.

Figure 6-6 shows an example in which factor levels are defined through visualization of a CDF. Generally, defining factor levels through visualization of a CDF is a matter of practitioner judgment rather than a specific procedure. For example, a practitioner may distinguish three different intervals in the lag period CDF, as the CDF shows a change in the slope at the

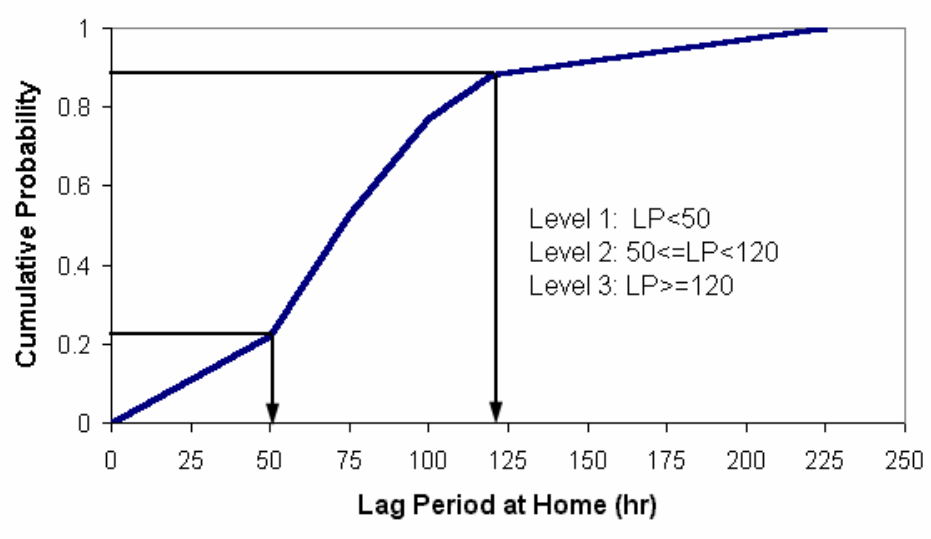

Figure 6-6. Definition of Levels for Lag Period Based Upon Visual Inspection of CDF.

boundaries of these intervals. These intervals correspond to the lower tail, middle range, and upper tail of the lag period variation domain. The levels may contain unequal sample size. Here, the middle level contains approximately 70 percent of generated values.

A key criterion for selecting number of levels for each input is with respect to the number of data points that lie within factor levels. The number of data points in each factor level directly affects the power of statistical tests for the analysis. Hence, there is always a trade-off between selecting higher number of factor levels and getting statistically significant results from ANOVA. However, there is also a trade-off between the desired number of iterations (e.g., in a Monte Carlo simulation), that are used to populate factor levels, and the computational time.

The methods suggested here for defining factor levels have some advantages and disadvantages. Defining levels with equal intervals helps identify possible thresholds in the model response using contrasts. With an increase in the number of levels for each factor and the use of similar intervals, thresholds can be estimated with more accuracy. A trade-off of this approach is the decrease in the number of data points within each level as the number of levels increases leading to lack of statistical power for estimated statistics. Using equal percentiles for definition of levels guarantees an equal number of data points in each level, thereby leading to a balanced experiment, which simplifies statistical analyses. Using visualization of the CDF for defining factor levels facilitates the evaluation of the model response in the lower or upper tail of

an input distribution. For example, ANOVA can compare the mean output values for lower and upper tail values of an input, if lower tail and upper tail are defined as separate factor levels.

After determining the factor levels, the next step is to choose the terms to include in the ANOVA model. The terms may include main effects of factors and interaction effects between factors. Based on the objectives of the analysis, a practitioner might choose only main effects in the ANOVA model for simplicity. Similar to regression analysis, the coefficient of determination, R^2, provides insight regarding whether the selected effects adequately capture variability in the output. A practitioner can use the R^2 value to determine whether additional terms are needed in the ANOVA model. A high R^2 value implies that a substantial amount of output variation is captured by the terms included in the ANOVA model. A high value of R^2 implies that results are not compromised by inappropriate definition of the levels for a factor. If the R^2 value is not satisfactory based upon the analyst's judgment, incorporation of more interaction terms between factors, adding higher-order interaction terms (e.g., three-way interactions), redefining factor levels, or transformation of the response variable may improve the R^2 value.

For two-dimensional probabilistic frameworks, it is preferable to use the same terms, including main effects of factors and interactions between factors, in each realization in order to facilitate the comparison of results among the realizations. Alternatively, an analyst can define factor levels and specify the effects to be investigated based upon analysis of one dataset selected at random (e.g., representing one realization of uncertainty), and use these same factor levels and effects to investigate all other random samples (e.g., representing all other realizations of uncertainty) to arrive at distributions of R^2 values. The evaluation of the R^2 distribution can reveal whether selected factor levels are appropriate, and whether terms included in the ANOVA model sufficiently explain the output variability.

After the analyst is satisfied with the model, the next step is to apply ANOVA to perform sensitivity analysis. The results from ANOVA include F values for each effect and the R^2 value. The details for the calculations of these values in ANOVA are given in Frey et al. (2003). In ANOVA, F values are used as sensitivity indices for ranking factors (Carlucci, 1999). As an alternative approach, the amount of explained variability corresponding to a factor effect, including simple and interaction effects, can be used as a sensitivity index (Rose *et al.,* 1991).

For probabilistic inputs, the F value is estimated based upon random samples from the input distributions. Hence, the F value is itself a random variable.

When an analyst is satisfied with the ANOVA model, additional insights regarding the sensitivity of the output to individual variation of each factor or simultaneous variation of factors can be achieved using contrasts. A contrast should be formed for factors or interaction effects that have a statistically significant effect. For example, a contrast can be built to evaluate the difference in mean pathogen growth when the storage temperature varies between high and low levels for a specific storage time.

A key question in using F values as sensitivity indices is regarding how much the F values of two factors must differ in order to discriminate their relative importance. Each simulation of a model may include multiple variability iterations and/or uncertainty realizations. Each simulation of a model obtains an estimated set of F values for factors. Therefore, simulation modeling can be used to estimate the sampling distribution of the F statistic. When compared with uncertainty in the F values for other inputs, it is possible to infer whether the rankings of the sensitivity of each factor would be unambiguous.

In some applications, it maybe impractical to run the model for hundreds of simulations as each simulation may take a few hours. Alternatively, bootstrap simulation can be used to generate sampling distributions of uncertainty for F values. Bootstrap simulation is a numerical method for estimating sampling distributions of statistics (Efron and Tibshirani, 1993). There are several variants of bootstrap simulation such as empirical bootstrap and parametric bootstrap. The empirical bootstrap method is suggested for using in estimating uncertainty of an F value. In the empirical bootstrap approach, an alternative randomized version of the original Monte Carlo simulation is obtained by sampling with replacement from the original set of random values. This procedure is computationally faster than generating new random samples from the original specified probability distributions of the model inputs (Efron and Tabshirani, 1993; Frey and Rhodes, 1998). ANOVA is applied to each of the bootstrap samples to produce a distribution of F values. A case study example is provided by Frey *et al* (2003).

As the final step of using ANOVA, estimated F values for the main effects of factors and interaction effects should be reported. Furthermore, the R^2 value or a distribution of R^2 values should be reported as a diagnostic check. If contrasts are used for further analysis of output variation, results of such analyses should also be included in the final report.

6.2.4 CART

This section introduces a procedure for application of CART for sensitivity analysis. CART is not included in available commercial statistical software packages; hence, the procedure presented here is based upon the S-PLUS™ statistical software package that was used by Frey *et al.* (2003). A key point in the procedure is that it relies heavily on visual recognition of patterns in the regression tree. As a result, CART is typically favored for application to one-dimensional probabilistic simulations. Application of CART for two-dimensional probabilistic simulations may be impracticable, if visualization of the regression tree is relied on to assess importance. However, the automation of CART can be facilitated by defining an alternative quantitative sensitivity index (e.g., some of deviance reduction associated with each input) as described later. Figure 6-7 presents a schematic diagram of the procedure for using CART as a sensitivity analysis method.

Assuming that a dataset is available for sensitivity analysis, the first step for a practitioner in applying CART is to define the characteristics of inputs in the dataset. As an example, an input in the dataset may represent a continuous variable, categorical variable, or a logical variable that can only hold true or false values. At each node in a regression tree, CART selects an input and a value within the range of variation of the input to split the dataset into stratified data subsets. The selected split-point of an input forms a condition on the input that determines which data values are assigned to each stratum. The appropriate specification of the characteristics of each input ensures that the selected split-points at a given node are plausible. For example, if input *A* is a discrete input with integer values of 1, 2, and 3, then a split-point could be anyone of these values but not 1.3.

As shown in Figure 6-7, the next step is to specify the criteria that govern how many times the data will be stratified. These criteria include: (1) minimum number of available observations before the splitting the dataset; (2) minimum sample size in a terminal node; and (3) minimum node deviance. For example, specifying a value of 100 for the first option means that CART requires at least 100 data points before a stratum of data at an internal node can be further split. The second option specifies the minimum sample size in a terminal node, which follows the final split in a regression tree. For example, a value of 50 for the minimum terminal node

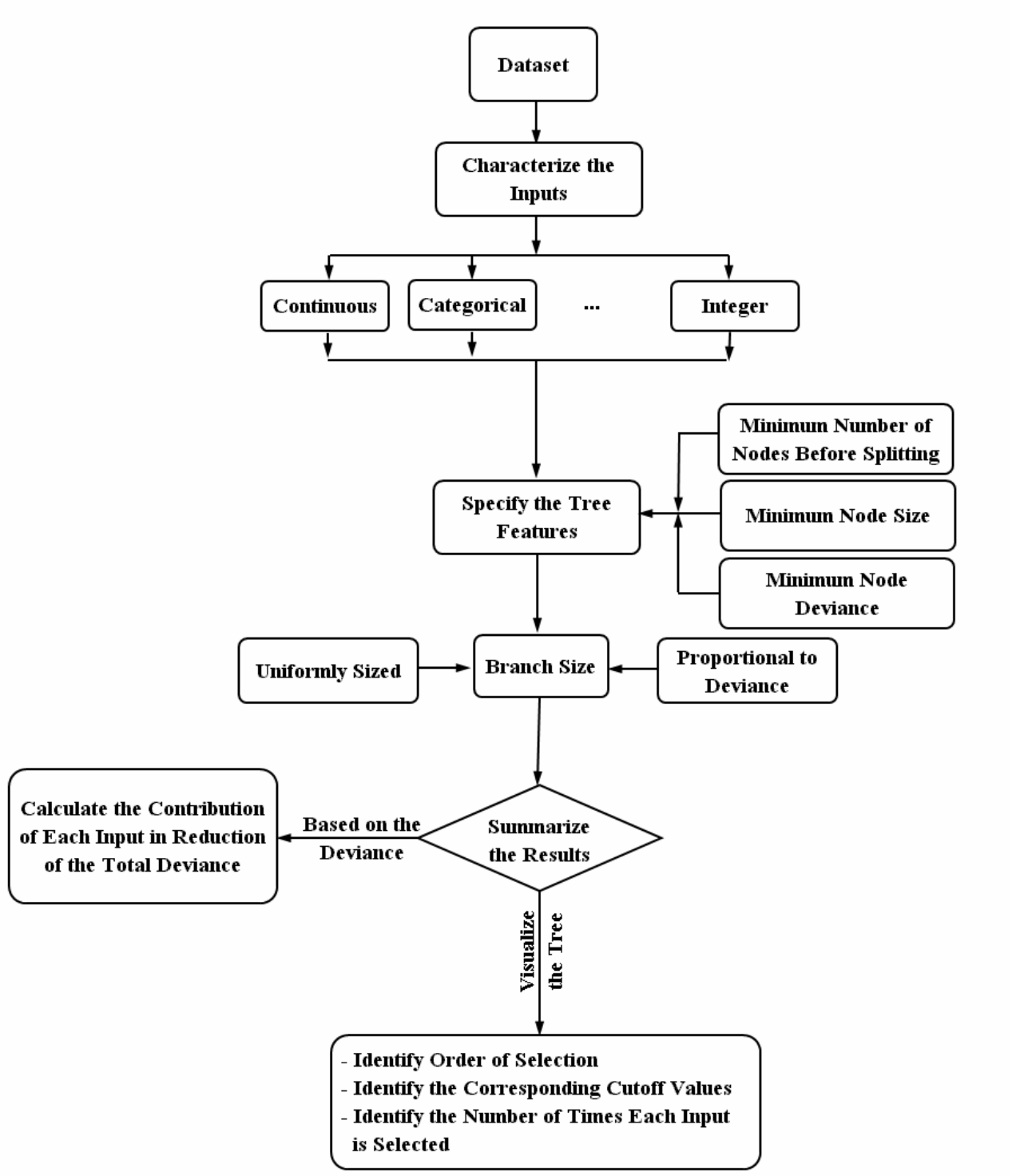

Figure 6-7. Schematic Diagram for Procedure of Application of CART.

sample size means that there must be at least 50 data points in each terminal node. The third option specifies a criterion with respect to the deviance of the stratified datasets. The deviance in each node represents the deviance of all data contained in the strata at a terminal or intermediate node. A minimum deviance of 0.1 for a terminal node indicates that the data would be split into stratified datasets such that terminal dataset has deviance of 0.1. The lower the minimum deviance, the further the dataset would be stratified.

The next step is to specify the branch size of the regression tree. The branch size represents the length of the vertical line that connects each node to its predecessor or successor node. This option determines the way that a regression tree is visualized. S-PLUS™ provides two options regarding the branch size of a tree: (1) proportional to the reduction in deviance; and (2) uniformly sized. When the former option is used, the length of the vertical lines in a regression tree between nodes will be displayed proportional to the reduction in deviance between two consecutive splits. This option should be selected for cases in which a practitioner not only wants to consider the order in which inputs are selected in a tree as a measure of sensitivity, but also wants to determine the contribution of each input to reduction of the total deviance over the entire tree. For example, if an input is selected at a node followed by a long vertical line before the next split, it indicates that splitting the dataset based on the selected input at the specified cut points substantially reduces the total deviance in the dataset. Further discussion regarding this issue is provided in Section 7.3.4. Regression trees displayed using branches proportional to the deviance require a large graphical display area when a large number of inputs are selected in the tree. In these cases, using branches of uniform length facilitates graphical display and visual inspection of the regression tree. If this option is selected for drawing a tree, a practitioner can only use the information regarding selected inputs and the order of selection in a tree to infer insights regarding sensitivity.

The last step in the procedure of application of CART is the summarization of the results. A practitioner can summarize the results of CART based on either visualization of the regression tree or by summarizing the contribution of each input to the reduction in the total deviance. If a practitioner relies on visualization, reported information include the order in which inputs are selected in a tree, the number of times that each input is selected in a tree, and classification rules provided in each node of a tree. A classification rule is a set of conditions that lead to a specific

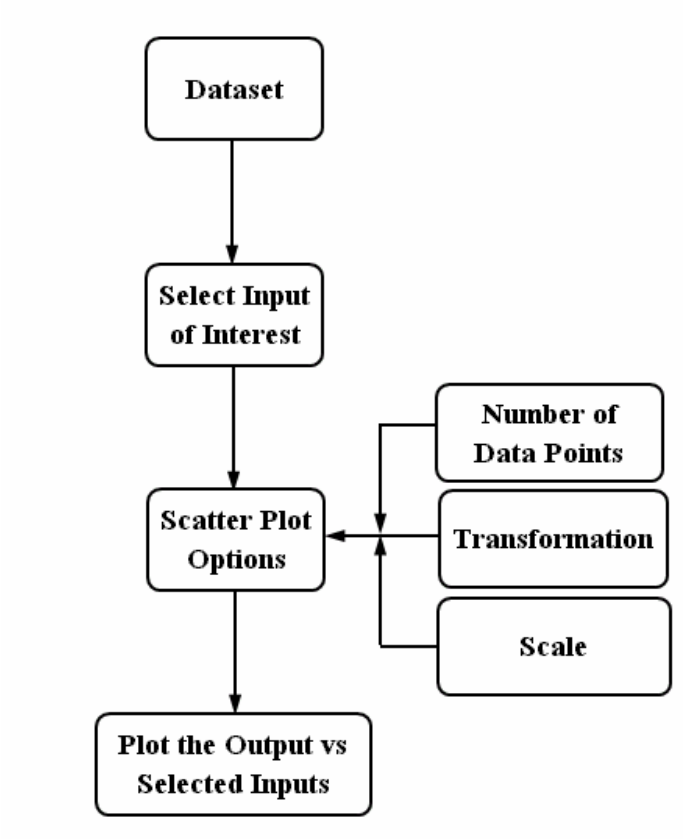

Figure 6-8. Schematic Diagram for Procedure of Application of Scatter Plots.

terminal node (i.e., the node with the highest mean exposure). S-PLUS™ provides an output file regarding the amount of deviance in the data at each node of a regression tree. A practitioner can use this information to estimate an alternative sensitivity index based upon the amount of contribution of each input to the reduction of the total deviance.

6.3 Procedure for Application of Graphical Sensitivity Analysis Methods

In this section, the procedure for application of graphical sensitivity analysis methods is presented. Two methods are discussed in this section including: (1) scatter plots; and (2) conditional sensitivity analysis. The procedures for application of the two methods are described in Sections 6.3.1 and 6.3.2, respectively. Each of the two methods has been briefly described in Section 5.3.3.

6.3.1 Procedure for Application of Scatter Plots

This section presents a procedure for application of scatter plots for sensitivity analysis. A key assumption for application of scatter plots is that a dataset including estimated output

values and sampled values from probability distributions of inputs is available. Figure 6-8 presents an application procedure for scatter plots.

Scatter plots can be presented in two or three dimensions, depicting variation of the output versus variation of one or two inputs. In order to prepare a scatter plot, as shown in Figure 6-8, the first step is to select the inputs of interest. A practitioner may select model inputs for based upon the results from other sensitivity analysis methods.

The next step in drawing the scatter plot is to specify the number of data points to display in the scatter plot. The number of data points displayed in a scatter plot should be dense enough to observe the appearance of any pattern, but not so dense that the variability within the scatter is difficult to observe (Vose, 2000). A practitioner may need to consider a transformation on output and input values before application of scatter plots. For example, a log transformation may be used to better reveal specific functional relationships between output and input values. The scale of the axes can influence the appearance of scatter plots to facilitate the interpretation and identification of possible patterns and relationships between the output and the selected inputs.

At the final step, a practitioner should graph output values versus the input values. Available software packages support the analysis of multivariate cases using matrix scatter plots (Saltelli, 2000). A matrix scatter plot is a rectangular array of scatter plots. Each element of the matrix is an individual scatter plot. Using this technique, hundreds of plots can be generated quickly and easily with an appropriate software.

6.3.2 Procedure for Application of Conditional Sensitivity Analysis

This section presents a procedure for application of conditional sensitivity analysis. As shown in Figure 6-9, the first step in the application procedure of conditional sensitivity analysis is to identify nominal values for the model inputs. Nominal values for the model inputs can be identified using the approach explained in Section 6.1 for NRSA.

The next step in the application procedure is to select an input for conditional sensitivity analysis. A practitioner may select model inputs based on the results from other sensitivity analysis methods.

After determining the input of interest, the next step is to perform a probabilistic simulation of the model in which samples are drawn from the probability distribution of the selected input, while other model inputs are fixed at their nominal values (e.g., minimum, mean, or maximum).

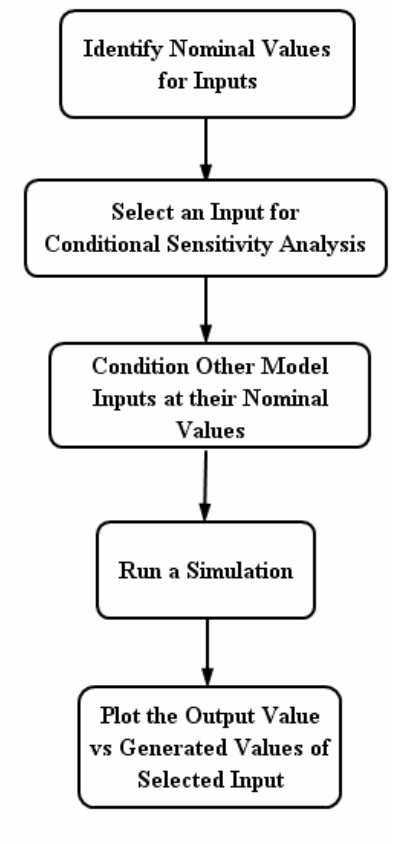

Figure 6-9. Schematic Diagram for Procedure of Application of Conditional Sensitivity Analysis.

At the next step, a practitioner should prepare a scatter or line plot that shows the variation of the output values versus generated values of the selected input in each of several cases, such as when other inputs are fixed at minimum, mean (or median), and maximum values.

6.4 Summary

This chapter introduced the application procedures for selected sensitivity analysis methods that were briefly described in Chapter 5. Each application procedure was presented in the form of a flowchart showing the key steps that a practitioner should follow when applying a sensitivity analysis method.

The presentation and interpretation of sensitivity analysis results are another important issue in sensitivity analysis. The next chapter describes how to present and interpret sensitivity analysis results for each selected method.

7 PRESENTATION AND INTERPRETATION OF RESULTS

Presentation and interpretation of the sensitivity analysis results are of high importance. An analyst should have good understanding of the theory underlying specific sensitivity analysis method in order to correctly interpret results. Good presentation of the results facilitates the interpretation step. A misinterpretation of the results may lead to ineffective decisions regarding risk management and control strategies. This chapter describes effective presentation and interpretation of results for each of the selected methods discussed in Chapter 6.

Section 7.1 discusses general principles for presenting and interpreting the sensitivity analysis results. Sections 7.2 through 7.4 discuss the interpretation and presentation of results based upon mathematical, statistical, and graphical sensitivity analysis methods, respectively. Each section provides examples from case scenarios presented by Frey *et al* (2003). Section 7.5 is a summary for this chapter.

7.1 General Principles in Presentation and Interpretation of Sensitivity Analysis Results

The general principles for presentation and interpretation of sensitivity analysis results include: (1) clearly identify the target audience, and tailor the presentation to that audience; (2) explicitly convey the objectives of the analysis; (3) describe the scenario so that the scope of the analysis is clear; (4) describe the model at an appropriate level of detail for the target audience; (5) describe the rationale for the selected sensitivity analysis methods; and (6) clearly present the results of the sensitivity analysis. Each of these general principles is briefly described in Sections 7.1.1 to 7.1.6, respectively.

7.1.1 Identify the Target Audience

The purpose of communicating the sensitivity analysis results to an audience is to ensure the proper and effective exchange of information and insights from the results of the analysis. The audience may include analysts, stakeholders, media, decision-makers, or some combination of these. The audience may include representatives of government agencies, private companies, non-profit organizations, academic institutions, or other organizations. Furthermore, the audience may include persons with a range of academic backgrounds and experience, including (for example) liberal arts, food science, microbiology, toxicology, sociology, economics, medicine, engineering, law, or others. Thus, there may be substantial variability in the background and perspectives among members of the intended audience.

The methods for presenting and interpreting results may be different for different audiences. Therefore, a basic principle for communicating sensitivity analysis results is that the audience should be clearly identified. For example, when presenting the sensitivity results to an audience with a limited scientific background, extensive information regarding the technical details of the methodology is probably not helpful or could be counter-productive. The time available with such an audience could be better spent conveying the objectives, scenarios, general characteristics of the model, general approach used for sensitivity analysis, and presentation of the sensitivity analysis results in plain language with the use of some supporting graphics. Qualitative information may be more useful than quantitative information for such audiences. If the objective is to provide insight into priorities for CCP and the selection of critical control limits, then the presentation could convey the relative importance of each possible CCP in either qualitative terms or using a simplified sensitivity index.

If the target audience is primarily other analysts, or persons with a strong technical background, then it is necessary to provide the most important technical details regarding the methods used, the models used, input assumptions, and interpretation of results, in addition to the basic qualitative information regarding objectives and scenarios. Moreover, the results of the analysis must be reproducible. However, the use of jargon should be with care. Even for a technical audience, it is likely that not everyone is familiar with the same jargon because each person may have a different technical background. Thus, any terminology unique to sensitivity analysis should either be avoided or will need to be clearly defined for the benefit of the audience.

7.1.2 Convey the Objectives of the Analysis

An analyst should ensure that the primary objectives of the sensitivity analysis are clearly communicated to the target audience. As a rhetorical device, it is often helpful to state the objectives in the form of key questions that will be answered by the results of an analysis. For example, one can state that the objective of the analysis is to answer one or more of the following key questions:

- Which controllable model inputs contribute the most to variation in exposure and risk?
- Which of the potential critical control points are the highest priorities in terms of avoiding the highest exposures?

- What critical limits are likely to achieve the specified risk management objectives?
- What are the key sources of uncertainty in the analysis?
- Which uncertainties are the highest priorities for reduction based upon additional data analysis or research?
- How well or appropriately does the model respond to changes in input values?

7.1.3 Describe the Scenario of the Analysis

As described in Chapter 4, a scenario is a set of assumptions about the nature of the problem to be analyzed. It is necessary to clearly describe the scenario of the sensitivity analysis when presenting the results to audience since the results may be specific to the particular assumptions made in a scenario. Clear introduction to the scenario will help the audience understand the implications of the results. For example, audience may determine whether the insights from the sensitivity analysis results can be generalized.

The presentation of a scenario should include the following components: (1) pathogens and populations analyzed; (2) pathways of interest identified and selected food categories; (3) spatial and temporal dimensions of the model simulation; and (4) probabilistic features. The details for the components are given in Chapter 4.

7.1.4 Describe the Model Used for the Analysis

Introductions to a model used for the sensitivity analysis are informative for the audience. Summarization of the key characteristics of a model can help the audience better understand the reasons for application of the selected sensitivity analysis methods and evaluate the credibility of the sensitivity analysis results. The presentation of the model used for sensitivity analysis should include: (1) model characteristics; and (2) model inputs and the output of interest. Model characteristics include linearity, thresholds, interactions, and information regarding the probabilistic framework of the model. Probability distributions of the inputs should be provided including specification of whether the input represents variability, uncertainty, or both. Model outputs should be clearly presented to the audience.

7.1.5 Describe the Rationale for the Selection of Sensitivity Analysis Methods

Selection of particular sensitivity analysis methods depends on model characteristics, available resources, or objectives of a sensitivity analysis. Presentation of the rationale for selection of the sensitivity analysis method(s) will help to establish the credibility, relevance, and

validity of the analysis. For example, when a model is non-linear and non-monotonic, the analyst can justify why ANOVA was selected while other sensitivity analysis methods such as correlation analysis were not used. Chapter 5 provides further discussion regarding the considerations in sensitivity analysis method selection.

7.1.6 Presenting Sensitivity Analysis Results

Sensitivity analysis results can be expressed in a variety of forms, such as using point estimates or average ranks of inputs, probability distributions of ranks under uncertainty, and graphs. When several sensitivity analysis methods are applied to a model, the results may be presented in different forms. Understanding of the results of sensitivity analysis by an audience strongly depends on a clear and informative presentation of the results. Tables and graphs are two common ways to present the sensitivity analysis results. Tables summarize the results of sensitivity analysis, while graphs can give the audience visual insights regarding the sensitivity analysis results.

Sections 7.2 to 7.4 discuss the presentation of the sensitivity analysis results using tables and graphs for selected methods.

7.2 Mathematical Sensitivity Analysis Methods

This section discusses the presentation and interpretation of sensitivity analysis results based upon NRSA and DSA. A brief description of the two methods is presented in Section 5.3.1. Because the presentation and interpretation of results from the two methods are similar, the discussion below is applicable to both methods.

Sensitivity analysis results from NRSA and DSA are expressed as sensitivity indices. Sensitivity indices are a set of numerical values used to indicate the magnitude of sensitivity of a given model output to each model input. The calculation of sensitivity indices based upon the two methods is given in Frey et al. (2003). The sensitivity indices can be negative or positive values based upon the model responses at nominal values of the model inputs. Model inputs can be prioritized or ranked based upon the relative magnitude of the sensitivity indices.

Two options are suggested for use in presenting mathematical sensitivity analysis results. One is to summarize the sensitivity indices and corresponding ranks for model inputs, as shown in the Table 7-1. Another is to use tornado graphs as shown in Figure 7-1. The summary table, graph, and the interpretation of the results are based on an example of the application of NRSA to the exposure part of the *Listeria* model (Frey, *et al.,* 2003).

Table 7-1. Results of Application of Nominal Range Sensitivity Analysis to the *Listeria monocytogenes* Exposure Module for Deli Salad (Source: Frey *et al.,* 2003)

Model Inputs	Nominal Range Sensitivity Index	Rank Within the Food Category
Serving Size (g)	71.0	2
Initial LM Concentration (log cfu/g)	94.7	1
Storage temperature (^{0}C)	31.7	3
Storage time (days)	7.1	5
Growth at 5 ^{0}C (log cfu/day)	13.5	4

As shown in Table 7-1, the summary table has columns of *names of model inputs*, *nominal range sensitivity index*, and *corresponding ranks*. The column of *names of mode inputs* lists inputs of interest; The column of *nominal range sensitivity index* shows the calculated sensitivity indices for the corresponding model inputs. The column of *corresponding ranks* display the rank order of the sensitivity of a given model output to each of the model inputs. A rank of 1 is considered to be the highest rank. A higher rank order indicates that the corresponding model inputs have larger individual contributions to variation in model output. For example, the input "Initial LM Concentration" has a rank of "1", implying that it is the most important input.

Tornado graphs are similar to bar graphs in which the abscissa holds sensitivity index values, while the ordinate shows inputs. In tornado graphs, a bar with a larger absolue value index indicates that the corresponding model input has a larger contribution to the variation in the model output. For example, in Figure 7-1, input *B* has the highest index value among all inputs to the model output, while the output shows lowest sensitivity to input *C*.

Both options introduced above can effectively communicate sensitivity analysis results from mathematical methods to the audience. Selection of either of them or both depends on the type of the audience. For the general public, a graphical presentation may be the most appropriate.

Because DSA does not consider the full range of variation in an input, the results may differ from those obtained from NRSA. Therefore, a comparison of results of DSA and NRSA can provide insight into whether the range of variation in the input contributes to sensitivity of the output to individual inputs.

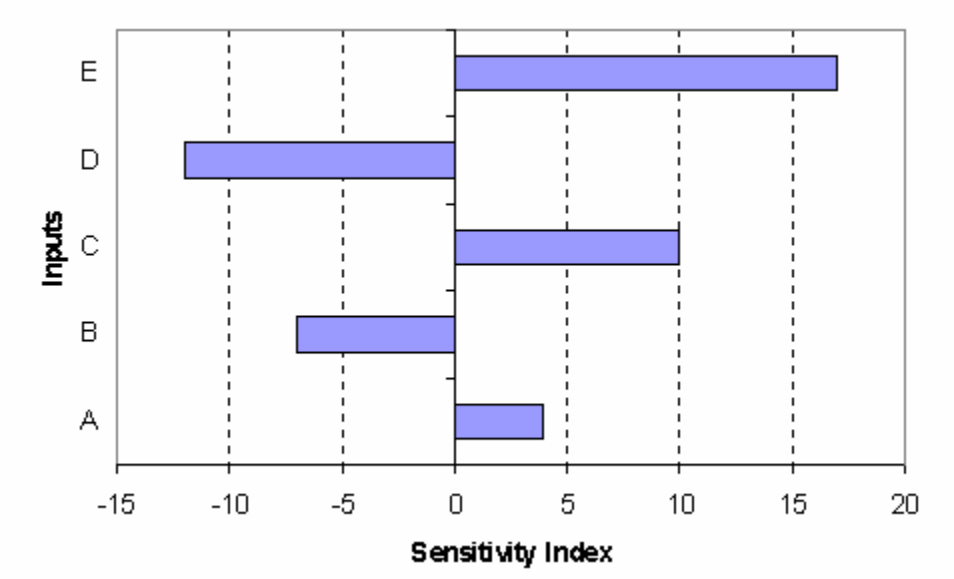

Figure 7-1. Tornado Graph for the Results of NRSA.

7.3 Statistical Sensitivity Analysis Methods

This section discusses the presentation and interpretation of sensitivity analysis results based upon statistical methods. A key feature of statistical sensitivity analysis methods is that they often involve one or two-dimensional probabilistic simulation of a model. The presentation and interpretation of the results may differ depending on the simulation approach and sensitivity analysis method. This section illustrates the presentation and interpretation of results for alternative probabilistic simulations of a model for the statistical methods discussed in Chapter 5. These statistical methods include sample and rank correlation coefficient methods, sample and rank regression analysis, ANOVA, and CART. Details regarding the presentation and interpretation of the results from these methods are provided in Sections 7.3.1 through 7.3.4, respectively.

7.3.1 Correlation Analysis

Correlation analysis measures the strength of a linear or monotonic relationship between an output and each of the model inputs, depending upon whether sample or rank correlations are used. Correlation coefficients have numerical values varying between -1 and 1. Larger absolute values of correlation coefficients indicate higher sensitivity of the output to the model input. Because the manner of presenting and interpretation of the results are similar for sample and rank correlation methods, the discussions below are applicable to both methods.

Table 7-2. Summary of the Pearson Correlation Coefficient Results for Two-Dimensional Variability Simulation for 100 Uncertainty Realizations (Source: Frey *et al.,* 2003)

Model Input [(a)]	Mean Correlation Coefficient	95% Probability Range of Coefficients	Frequency	Mean Rank	Range of Rank
Temperature 1	0.269	(0.030,0.525)	93	4.3	1-13
Temperature 2	0.008	(-0.070,0.102)	7	10.9	5-13
Temperature 3	0.466	(0.290,0.624)	100	1.6	1-6
Time1	0.252	(0.051,0.478)	95	4.7	1-10
Time2	-0.005	(-0.064,0.061)	2	10.9	8-13
Time3	0.339	(0.123,0.491)	98	3.9	1-13
Maximum Density	0.027	(-0.052,0.095)	10	10.6	6-13
Lag Period 1	-0.169	(-0.317,-0.022)	83	7.0	3-13
Lag Period 2	-0.009	(-0.104,0.071)	7	10.8	5-13
Lag Period 3	-0.311	(-0.45,-0.146)	100	4.8	2-8
Generation Time 1	-0.168	(-0.328,0.00)	79	7.1	2-13
Generation Time 2	-0.006	(-0.092,0.065)	7	10.7	6-13
Generation Time 3	-0.339	(-0.496,-0.162)	100	3.8	2-8

(a) 1 = Stage 1 (i.e., retail); 2 = Stage 2 (i.e., transportation); and 3 = Stage 3 (i.e., home).

Tables and graphs can be used for presentation of the results from sample and rank correlation coefficient methods. However, based upon the type of probabilistic simulation, the contents of the table may differ. For example, for one-dimensional analyses, the table used for summarizing the results could be similar to Table 7-1 used for mathematical methods, except that there should be additional information to indicate whether each estimated coefficient is statistically significant. Insignificant inputs can be assigned the lowest ranks, indicating no measurable sensitivity. For two-dimensional cases, the table should present a statistical summary of the results for all of the uncertainty realizations or variability iterations, as in Table 7-2, which summarizes the results of an application of sample correlation to the growth estimation part of the *E. coli* model for a two-dimensional probabilistic simulation of 650 variability iterations under 100 uncertainty realizations of the model (Frey *et. al.,* 2003).

As shown in Table 7-2, there are six columns. The column of *Model Input* lists the model inputs of interest. The column of *Mean Correlation Coefficient* shows the average of *n* correlation coefficients for the corresponding model inputs, which are calculated from *n* uncertainty realizations. In this case, *n* is 100. The column of *95% Probability Range of Coefficients* displays the variation in coefficients under *n* uncertainty realization for a 95% probability range. The *Frequency* column represents the number of the correlation coefficients

that are statistically significant under *n* uncertainty realizations. For example, for *Temperature 1*, there are 97 times out of 100 uncertainty realizations that the calculated coefficient was statistically significant. The column of *Mean Rank* is the arithmetic mean of the ranks in each of the uncertainty realizations. This column indicates the overall sensitivity of the model output to a particular model input. The column of *Range of Rank* summarizes the variation in ranks for a model input among the *n* uncertainty realizations and provides insight regarding the degree of ambiguity in relative importance of an input based upon uncertainty in the input.

Summary results provided in Table 7-2 can facilitate identifying groups of model inputs with similar sensitivity. For example, the model inputs in this example can be approximately classified into four groups. The four groups are: (1) the most important input, storage temperature at stage 3, which has a mean rank of 1.6; (2) inputs of secondary importance, including storage time at stages 1 and 3, storage temperature at stage 1, and generation time and lag period at stage 3, which have similar mean ranks of 3.9 to 4.7; (3) inputs of tertiary importance, including lag period and generation time at stage 1, with mean ranks of 7.0 and 7.1, respectively; and (4) inputs of minor or no importance, including inputs corresponding to the second stage (i.e., transportation) and the maximum density, with mean ranks of 10.6 to 10.9. Each of these groups has mean ranks that are distinguishable from the mean ranks in other groups.

Two types of graphs can be used to present the two-dimensional sensitivity analysis results from the correlation coefficient method. One is tornado graphs, as shown in Section 7.2, and another is complementary cumulative distribution functions (CCDF) of ranks.

Figure 7-2 presents an example of a tornado graph for the four most important inputs based upon the results given in Table 7-2. Compared to Figure 7-1 used to present sensitivity indices for mathematical methods, additional information is shown. Besides the mean values of coefficients, the 95 percent probability range for each coefficient is also shown in the graph. This presentation will help the audience understand the uncertainty associated with each correlation coefficient. Overlaps in the ranges of variation of multiple inputs indicate that there may be groups of inputs with similar importance.

Figure 7-3 shows an example of a CCDF graph of ranks for the four inputs with the highest sensitivity based upon the results given in Table 7-2. The figure shows the variation in

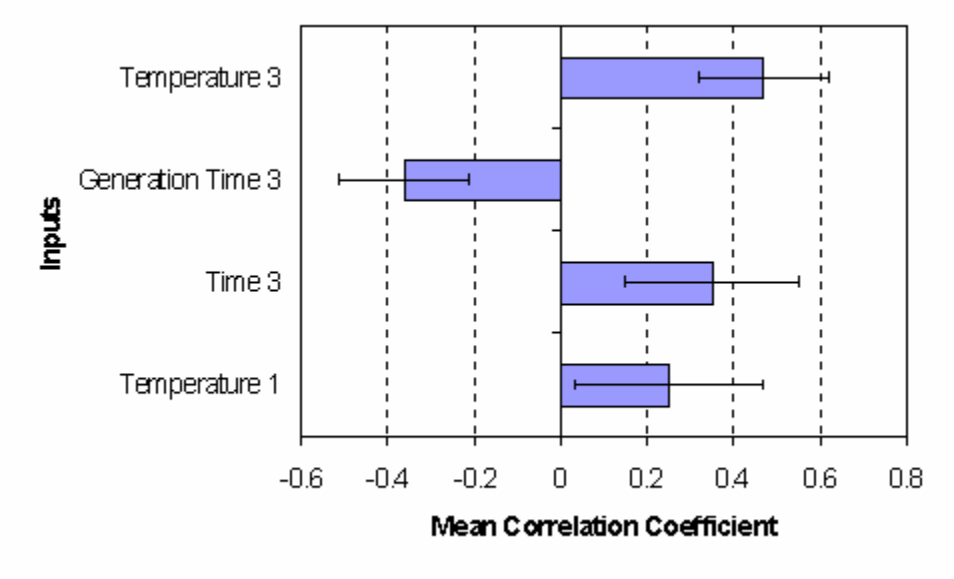

Figure 7-2. Tornado Graph for the Top Four Important Inputs Based on the Results of Sample Correlation Coefficient Method.

ranks of an input and the probability associated with the relative rank of the input in each uncertainty realization. The CCDF for the storage temperature at stage 3 (Temperature 3) indicates that the rank was equal to one in 65 percent of the simulations, which implies that the rank was worse than one for 35 percent of the simulations. Furthermore, the storage temperature at stage 3 was ranked sixth or higher for 100 percent of the simulations. In contrast, generation time at stage 3 (Time 3)was never selected as the most important input in the uncertainty realizations of the model, while there was 100 percent probability of allocation of a rank better than eight to this input.

Figure 7-3 implies that the output has approximately the same sensitivity to three inputs of generation time at stage 3, storage time at stage 3, and storage temperature at stage 1. Note that the CCDFs for these inputs overlap to some extent and show a similar pattern. Although there is some ambiguity in the relative rank of storage temperature at stage 3, the obvious distinction of the CCDF for this input compared to those of other inputs implies that the output has a substantially higher sensitivity to variation of this input. The ranking of storage temperature at stage 3 as the most important input is robust to uncertainty.

7.3.2 Regression Analysis

This section discuses the presentation and interpretation of the sensitivity analysis results using sample and rank regression analysis methods. In both regression methods, the estimated

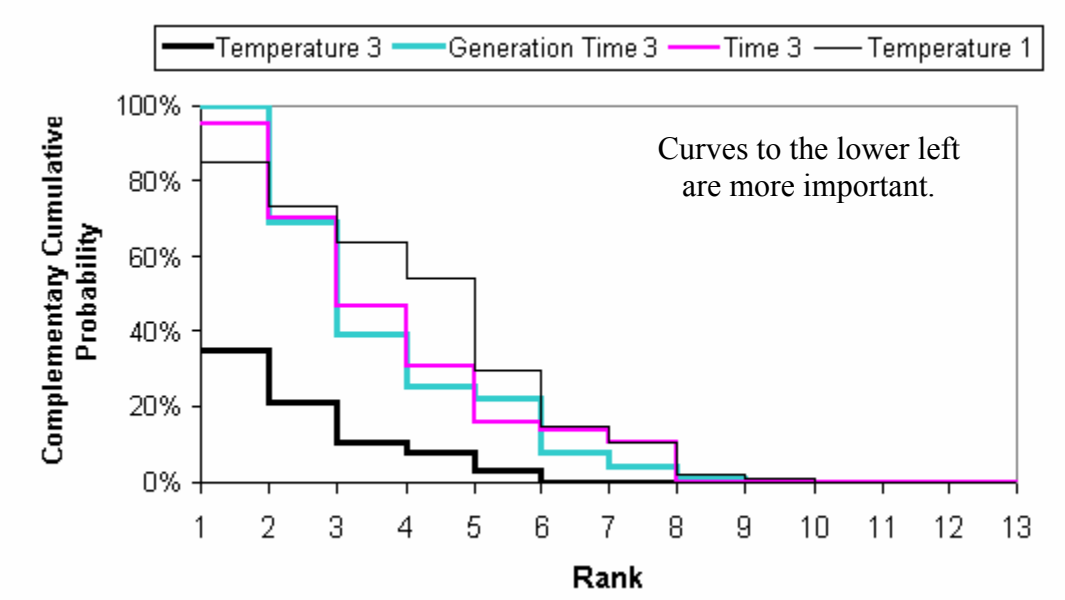

Figure 7-3. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected Inputs Based Upon Pearson Correlation: Storage Temperature at Stages 1and 3; Storage Time at Stage 3; and Generation Time at Stage 3.

regression coefficients are used as sensitivity indices to indicate the sensitivity of the output to the model inputs. For standardized sample regression analysis, the estimated coefficients are referred to as standardized regression coefficients. The calculation of the regression coefficients and more detail regarding regression analysis methods are given in Frey et al. (2003). Because the manner of presenting and interpretation of the results are similar for sample and rank regression analysis methods, the following discussions are applicable to both methods.

The validity of insights from sensitivity analysis depends upon the appropriateness of the regression model formulation. Model inputs can be ranked based upon the relative magnitude of estimated coefficients. Statistically significant coefficients indicate that there is at least some degree of sensitivity. Statistically insignificant coefficients indicate that the output is not sensitive to the corresponding inputs.

The statistical significance of coefficients for model inputs can be evaluated using *t* tests and corresponding *P* values. The *t* test provides a statistic to evaluate whether the estimated coefficient differs significantly from zero. For a regression coefficient to be judged statistically significant, the conventional cut-off for the P value is 0.05 (Neter *et. al.,* 1989).

Similar to correlation coefficient methods, both tables and graphs can be used to present the sensitivity analysis results for regression analysis methods. However, based upon the type of

Table 7-3. Sample Standardized Regression Analysis Results for the Growth Estimation Part Based Upon Variability Only ($R^2 = 0.51$)

Model Input [(a)]	Coefficient	95% CI [(b)]	*t* Value	Pr>F	Rank [(c)]
Storage Temperature, Stage 1	0.32	(0.31,0.32)	74	<0.0001	3
Storage Temperature, Stage 2	3×10^{-3}	$(-7,13)\times10^{-3}$	0.6	0.6	---
Storage Temperature, Stage 3	0.59	(0.58,0.60)	124	<0.0001	1
Storage Time, Stage 1	0.27	(0.26,0.27)	97	<0.0001	4
Storage Time, Stage 2	7×10^{-3}	$(2,13)\times10^{-3}$	3	0.1	---
Storage Time, Stage 3	0.34	0.34[(d)]	123	<0.0001	2
Maximum Density	0.012	(0.007,0.02)	5	<0.0001	7
Lag Period, Stage 1	-0.012	(-0.019, -0.01)	-3	0.0005	7
Lag Period, Stage 2	-1×10^{-4}	(-0.01,0.01)	0.0	0.3	---
Lag Period, Stage 3	-1×10^{-3}	(-0.01,0.01)	-1	0.4	---
Generation Time, Stage 1	0.08	(0.07,0.09)	20	<0.0001	6
Generation Time, Stage 2	-4×10^{-3}	(-0.014,0.01)	-1	0.7	---
Generation Time, Stage 3	0.11	(0.10,0.12)	23	<0.0001	5

(a) 1 = Stage 1 (i.e., retail); 2 = Stage 2 (i.e., transportation); and 3 = Stage 3 (i.e., home).
(b) CI = Confidence Interval for the coefficient.
(c) Statistically insignificant inputs are less important than the ranked inputs and are equivalent to a rank of 8 or more.
(d) The interval for this coefficient is so tight that it appears as 0.34 to 0.34 when it is rounded to two decimal places.

probabilistic simulation, the contents of the table or graph may be different. Typically, the table used for presenting a two-dimensional simulation results would contain information similar to the two-dimensional correlation analysis. This section will focus on the discussion of one-dimensional sensitivity analysis results. The presentation and interpretation of sensitivity analysis results using regression analysis methods are illustrated below with the help of an example of application of standardized regression analysis to the growth estimation pat of the *E. coli* model (Frey *et. al.,* 2003).

Table 7-3 shows an example summary table of the results using sample regression methods to analyze the growth estimation part of the *E. coli* model. The example is a one-dimensional simulation with variability only while keeping uncertainty parameters at mean values. Table 7-3 is comprised of six columns. The column of *Model Input* lists the model inputs of interest. The column of *Coefficient* displays the estimated regression coefficients for corresponding model inputs. The column of *95% CI* shows the 95% confidence intervals of the corresponding regression coefficient, which is calculated from the standard error of an estimated regression coefficient. Confidence intervals reveals ambiguity in ranks based upon the absolute value of the regression coefficient. Overlapping confidence intervals indicate no significant

difference in sensitivity to the model inputs. Ranks in the summary table reflect the relative sensitivity of the model output to model inputs. A higher rank shows a higher sensitivity of the model output to a model input. However, those model inputs that are not statistically significant are not ranked or could be assigned the lowest possible rank.

The coefficient of determination, R^2, can be used to assess whether the fitted regression model describing model variability is reasonable. A high R^2 value implies that the assumption for the functional relationship between the output and inputs is substantially valid. The R^2 value should be also reported in the summary table, and it can be put in the table title or notes. The R^2 value for the sample regression analysis was 0.51, indicating that approximately 50 percent of the output variation is addressed using the linear regression model. Although the variation captured by the regression model is not very high, it is still in the acceptable range, indicating the ranks provided by relative magnitude of regression coefficients may be reliable.

The rankings of the inputs in Table 7-3 based upon the regression coefficient values are generally unambiguous. For example, the top ranked input, storage temperature at stage 3, has a coefficient that is significantly larger than that of the second ranked input, the storage time at stage 3. The inputs ranked second through seventh are significantly different from each other in importance in that the confidence intervals for their coefficients do not overlap. Two inputs are both ranked seven because they have equal coefficients. These two inputs are maximum density and lag period at stage 1. The confidence intervals for these two inputs overlap. However, both inputs have coefficients that are substantially smaller than all other inputs. Therefore, the two inputs are of little importance compared to the other ranked inputs.

Bar charts can be used to present the one-dimensional results of the regression analysis including absolute values of regression coefficients and corresponding confidence intervals. Figure 7-4 graphically presents the results summarized in Table 7-3 for statistically significant inputs. The audience can easily gain insight that storage temperature at stage 3 is the most important input based upon the absolute magnitude of the standardized regression coefficients. Confidence intervals graphed in Figure 7-4 shows that the relative ranks of inputs are unambiguous, as those intervals do not overlap except for the two least important inputs.

To present the results from a two-dimensional analysis, the summary table is similar to Table 7-2. The summary table can include the names of inputs, mean regression coefficient for each input, 95 percent probability range for each coefficient, number of times that each

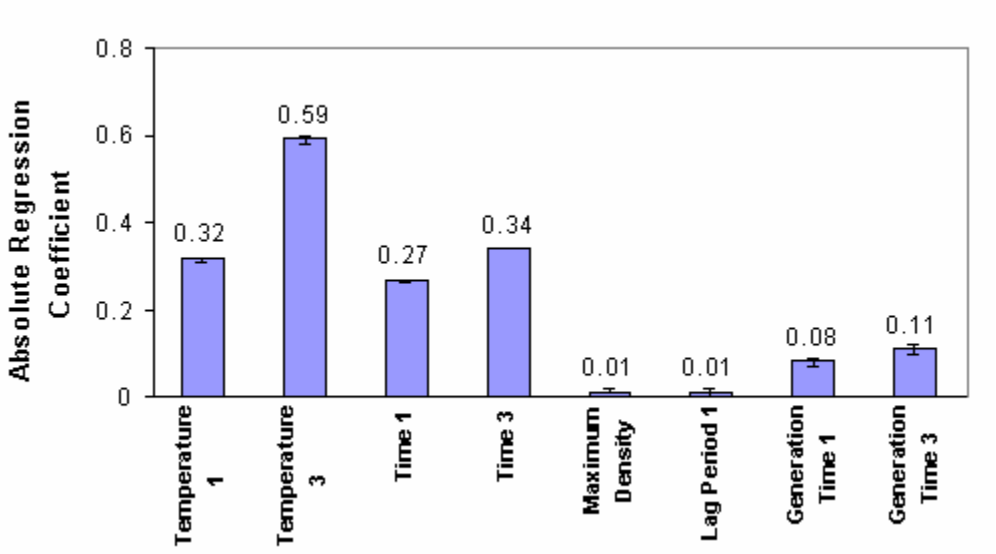

Figure 7-4. Example Bar Chart for Statistically Significant Inputs with Corresponding Confidence Intervals.

coefficient is identified as statistically significant, mean rank corresponding to each input, and range of ranks for individual inputs. CCDF graphs, as shown in Figure 7-3, can be used to present the results for regression analysis methods applied to a two-dimensional simulation. CCDFs provide insight regarding the range and likelihood of ranks corresponding to each input. CCDF graphs can facilitate identification groups of inputs with similar importance.

A CDF graph of the R^2 values can also be used to present the two-dimensional sensitivity analysis results when using regression methods. The CDF of R^2 provides insight regarding the uncertainty of goodness-of-fit in multiple realizations of a model. If both sample and rank regression analyses are applied to a model, comparison of the corresponding CDFs of R^2 values may reveal model characteristics such as non-linearity or monotonicity. Figure 7-5, for example, compares CDFs of R^2 values for sample and rank regression analyses applied to the slaughter module of the *E. coli* model. Figure 7-5 indicates that application of rank regression substantially improved the proportion of variability in the output accounted for by the regression model. Therefore, the slaughter module may have a monotonic, nonlinear association between the output and inputs.

7.3.3 Analysis of Variance

This section discusses the presentation and interpretation of the sensitivity analysis results using ANOVA. In ANOVA, the estimated F values for each factor and possible

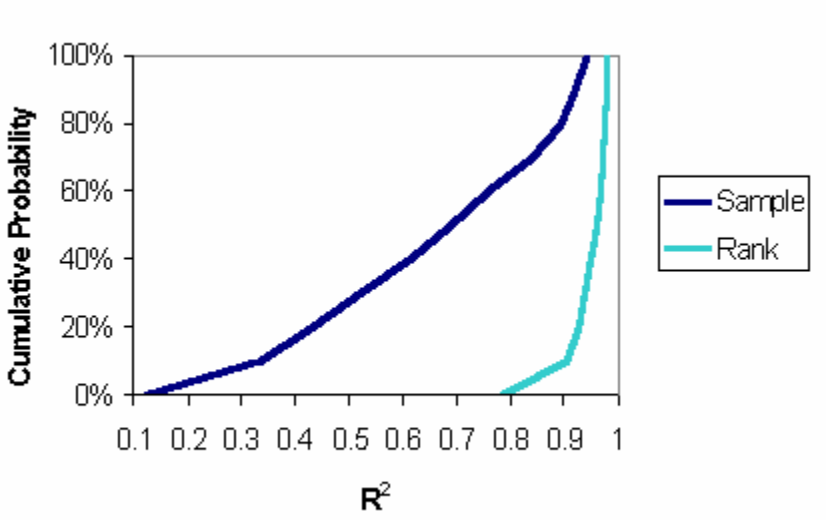

Figure 7-5. R^2 Distributions for Sample and Rank Regression Analyses.

interaction terms indicate the sensitivity of the model output to model inputs of interest. Details on the ANOVA method and procedure of implementing ANOVA are given in Sections 5.3.2.3 and 6.2.3, respectively.

F values estimated for each factor and possible interaction effects between factors are used as sensitivity indices. Factors and interaction terms are ranked based upon the relative magnitude of their F values. Similar to correlation and regression analysis methods, both summary tables and graphs can be used to present the results of ANOVA. The contents of the summary tables depend on the type of probabilistic simulation of the model. For one-dimensional scenarios, a summary table of ANOVA results should include estimated F values, corresponding P values, and relative ranks. Factors that are statistically significant (e.g., $P<0.05$) are ranked. When ANOVA is applied to a two-dimensional analysis of variability and uncertainty, the summary table is similar to Table 7-2 for Pearson correlation coefficient methods, except that mean F values are reported for each factor or interaction term instead of a mean correlation coefficient.

Bar charts representing F values can be used to convey insights regarding sensitivity. For two-dimensional scenarios, CCDF graphs similar to Figure 7-3 can be prepared to present the range and likelihood for each factor and the range of ambiguity in the associated ranks. These graphs can assist in identifying model inputs with similar importance.

Table 7-4. Summary of the ANOVA Results for 200 Bootstrap Simulations for F value Statistics

Model Factor [(a)]	Mean F Value	95% Probability Range	SD/Mean	Frequency (Percent)	Mean Rank	Range of Rank
Temperature, Stage 1	481	(318,606)	0.16	100	4.0	3-4
Temperature, Stage 2	1.8	(0.0,9.6)	1.45	17	10.9	7-13
Temperature, Stage 3	1010	(810,1180)	0.09	100	1.0	1-2
Storage Time, Stage 1	657	(557,780)	0.09	100	2.9	2-4
Storage Time, Stage 2	0.4	(0.0,2.6)	1.58	1	12.3	9-13
Storage Time, Stage 3	781	(714,915)	0.06	100	2.1	1-3
Maximum Density	8.6	(1.3,26)	0.71	79	8.1	7-13
Lag Period, Stage 1	50	(35,64)	0.14	100	5.9	5-6
Lag Period, Stage 2	1.5	(0.1,5.0)	0.76	15	10.6	9-13
Lag Period, Stage 3	60	(47,73)	0.11	100	5.1	5-6
Generation Time, Stage 1	16	(9.3,24)	0.26	100	7.2	7-8
Generation Time, Stage 2	1.7	(0.1,4.8)	0.74	17	9.9	8-12
Generation Time, Stage 3	19	(12,25)	0.21	100	6.8	6-8

(a) 1 = Stage 1 (i.e., retail); 2 = Stage 2 (i.e., transportation); and 3 = Stage 3 (i.e., home).

A key point in interpretation of the ANOVA results is that the ranking of factors based upon the relative magnitude of F values can be ambiguous. Substantial differences in F values of two factors indicate an obvious difference in sensitivity of the output to these factors. In order to quantify uncertainty or ambiguity in ranks based upon the magnitude of F values, the bootstrap technique can be used (Frey *et al.,* 2003). Results of the bootstrap simulation can be summarized in a table.

For example, Table 7-4 summarizes 200 bootstrap simulation results. ANOVA was applied to each bootstrap simulation result; therefore, 200 F values and ranks were estimated for each factor. The summary table is comprised of seven columns. The column of *Model Factor* lists the factors of interest in the model. The column of *Mean F Value* shows the average of F values for each model factor, calculated over *n* bootstrap simulations (in this case, *n* is 200). The column of *95% Probability Range* displays uncertainty in F values under *n* bootstrap simulation on 95% probability range. The column of *SD/Mean* presents the coefficient of variation. The *Frequency* column represents the number of the F values that are statistically significant over *n* bootstrap simulations for each factor. The column of *Mean Rank* is the arithmetic mean of *n* ranks from *n* bootstrap simulations. The column indicates the overall sensitivity of the model

output to a certain factor. The column of *Range of Rank* summarizes uncertainty in ranks for a factor under *n* bootstrap simulations and indicate the degree of ambiguity in relative importance of each factor based upon uncertainty in F values.

A practitioner can use the results in Table 7-4 to quantify the uncertainty in F values. For example, the storage temperature at stage 3 is estimated to have a mean rank of 1.0. The mean F value for this factor is 1,010, and the 95 percent probability range of the F value is 810 to 1,180, or approximately plus or minus 20 percent of the mean value. Storage time at stage 3 has a mean rank of 2.1, a mean F value of 781 and a 95 percent probability range of 714 to 915. The overlap in the confidence intervals for the two factors indicates tied rank order between the two factors can reverse, even though on average the F value for storage temperature is 1.3 times larger than the one for storage time. The F values for these factors do not have statistically significant correlation. In order to gain insight regarding how large the ratio of two F values must be in order for the ranks of the corresponding factors to be significantly different, the possible range of variation for the F values of each factor should be quantified.

For a specific case scenario by Frey *et al.,* (2003), statistically significant F values that differed by 30 percent or more were judged to be substantially different. These results are specific to the example provided here, however, should not be used to make general quantitative judgments regarding differences between F values obtained with different sample sizes or models. However, the case study result suggests that two factors having similar F values are probably of comparable importance.

A practitioner can use contrasts in ANOVA to obtain more insights regarding the sensitivity of the output to individual variation of each factor or simultaneous variation of multiple factors. A contrast is the comparison of the mean output values between two levels (or two sets of levels) of independent factors following ANOVA. Contrasts can be used to test complex patterns between mean output values. The issue of contrasts is discussed further in Frey *et al.* (2003) and Section 6.2.3 of this document. A practitioner can infer special model characteristics such as non-linearity, threshold and saturation points, and interactions by clear interpretation of the contrast results. Table 7-5 summarizes an example in which a set of contrasts is used to infer an interaction effect between two factors in a model and possible saturation points. This example is about the application of ANOVA to the growth estimation part of the *E. coli* model (Frey *et. al*., 2003).

Table 7-5. Evaluation of ANOVA Contrasts for the Growth Estimation Regarding the Interactions Between Storage Temperature and Storage Time at Stage 1

Contrast	Estimate (1)	F Value	Pr>F	Significant
T [7.5-11°C], Time 1st and 2nd days	0.005	112	<0.0001	Yes
T [7.5-11°C], Time 2nd and 3rd days	0.026	1280	<0.0001	Yes
T [7.5-11°C], Time 3rd and 4th days	0.049	1720	<0.0001	Yes
T [7.5-11°C], Time 4th and 5th days	0.069	1280	<0.0001	Yes
T [7.5-11°C], Time 5th and 6th days	0.074	610	<0.0001	Yes
T [7.5-11°C], Time 6th and 7th days	0.103	455	<0.0001	Yes
T [7.5-11°C], Time 7th and 8th days	0.042	30.8	<0.0001	Yes
T [7.5-11°C], Time 8th and 9th days	0.031	6.8	0.008	Yes
T [7.5-11°C], Time 9th and 10th days	0.108	38.8	<0.0001	Yes
T [7.5-11°C], Time 10th and 11th days	------	0.16	0.8	No
T [11-14.5°C], Time 1st and 2nd days	0.116	4060	<0.0001	Yes
T [11-14.5°C], Time 2nd and 3rd days	0.211	5290	<0.0001	Yes
T [11-14.5°C], Time 3rd and 4th days	0.218	2170	<0.0001	Yes
T [11-14.5°C], Time 4th and 5th days	0.119	240	<0.0001	Yes
T [11-14.5°C], Time 5th and 6th days	0.087	48.4	<0.0001	Yes
T [11-14.5°C], Time 6th and 7th days	------	0.9	0.6	No
T [18-21.5°C], Time 1st and 2nd days	0.55	2630	<0.0001	Yes
T [18-21.5°C], Time 2nd and 3rd days	-----	0.1	0.4	No
T [21.5-25°C], Time 1st and 2nd days	0.503	6270	<0.0001	Yes
T [21.5-25°C], Time 2nd and 3rd days	-----	2.4	0.09	No

(1) The *Estimate* column represents the estimate of the difference in the growth of the *E. coli* organisms in two consecutive days.

The summary table for the contrast results includes five columns. The *Contrast* column includes information regarding the specific levels of factors involved in the contrast. Storage temperature (T) and storage time (Time) at stage 1 are involved in the example contrasts. The *Estimate* column presents the estimated contrast. The columns of *F Value* and *Pr>F* list the calculated *F* and *P* values for the corresponding contrast, which are used to evaluate whether the estimated contrast is statistically significant. The column of *Significant* indicates whether the contrast is statistically significant. Contrasts with P values less than 0.05 are considered as statistically significant. If corresponding probability is larger than 0.05, there is not enough statistical support to indicate that the estimated value for the contrast is different from zero.

Contrast results in Table 7-5 indicate that when storage temperature in stage 1 is at the first level (i.e., between 7.5°C and 11°C), storage time influences growth of *E. coli* in ground beef until the tenth day. Statistically significant contrasts are estimated for the difference in pathogen growth for consecutive days through the tenth. After the tenth day there is no

significant difference between the estimated growth, indicating that maximum population density is achieved by the tenth day. When storage temperature at stage 1 is at the second level (e.g., between 11°C and 14.5°C), the contrasts indicate that maximum population density is reached in only five days. When the storage temperature at stage 1 increases to the third, forth and fifth levels, maximum population density is reached in four, three, and two days, respectively. This pattern indicates an interaction effect between storage time and temperature.

The coefficient of determination, R^2, should be reported as a diagnostic for the results of ANOVA. Although the F values calculated for each effect indicate the statistical significance of corresponding effect, the coefficient of determination indicates whether the selected effects adequately capture variability in the output. Generally, a high R^2 value implies that results are not compromised by incomplete specification of effects or by inappropriate definition of the levels for a factor. When ANOVA is applied to two-dimensional simulations of variability and uncertainty, a CDF graph for R^2 can provide information about uncertainty in the adequacy of the ANOVA model for multiple realizations of the model.

7.3.4 CART

This section discusses the presentation and interpretation of sensitivity analysis results using CART. CART was evaluated by Frey *et. al.* (2003) using several case studies based upon two food safety risk assessment models using the S-PLUS™ statistical software package. Presentation of CART results may depend on the software package used. Details regarding the methodology and procedure of application of CART are given in Frey et al. (2003) and Section 6.2.4, respectively.

In general, results of sensitivity analysis using CART are presented in the form of regression trees. CART reduces the total deviance in the dataset generated in the probabilistic simulation of the model by splitting the dataset into stratified datasets with more homogeneous variation of the output within each stratified dataset. The output is inferred to be most sensitive to those inputs selected in the regression tree, with the strength or importance of the dependence implied by the order, frequency, or both, with which an input appears in the tree, or based upon the role of an input with regard to classifying an outcome (e.g., high mean exposure) of interest.

Interpretation of the results based upon CART may not be straightforward, since CART does not produce a singular sensitivity index. The analyst has to scrutinize the regression tree via visualization of the tree and considering the order in which inputs are selected in the tree. Cut off

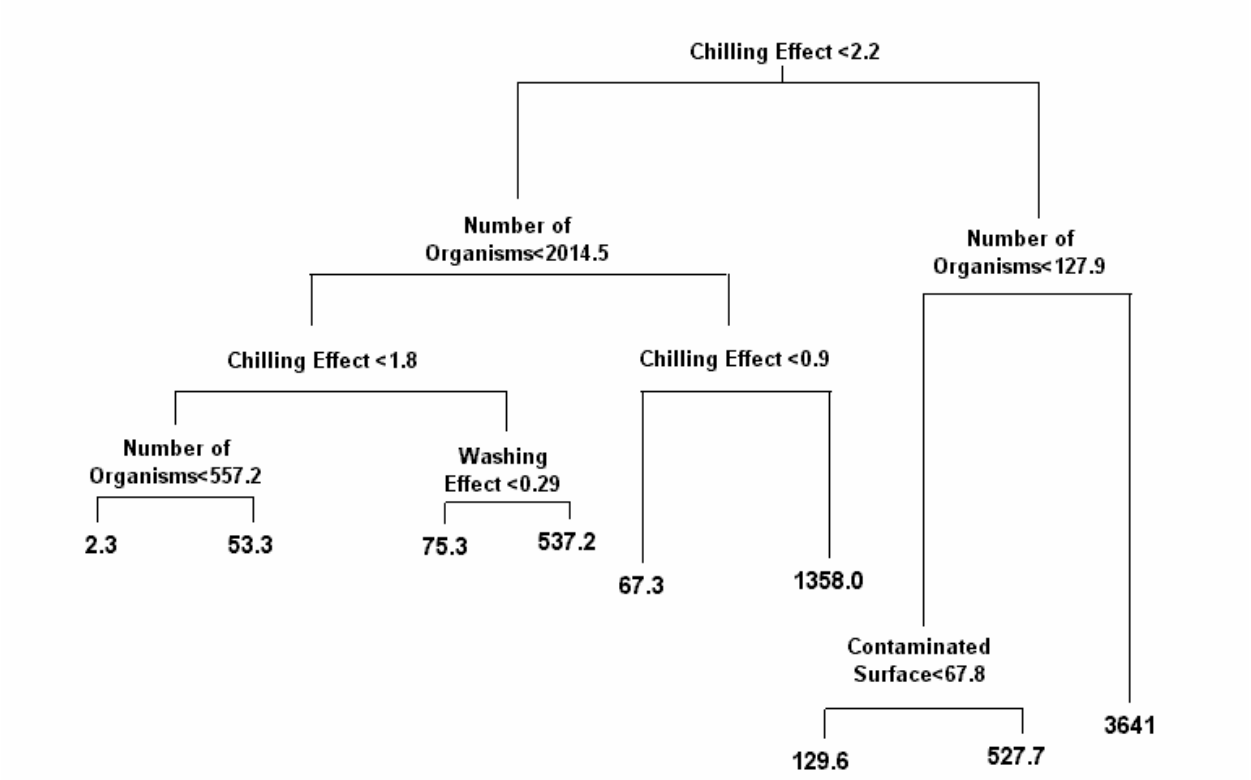

Figure 7-6. The Regression Tree for the Combo Bin Contamination from Steer and Heifer in Summer for One-Dimensional Variability and Uncertainty Analysis (Source: Frey *et al.,* 2003).

points selected corresponding to each input should also be considered for inferring special model characteristics, such as non-linearity or thresholds. As a general rule of thumb, the output typically has higher sensitivity to model inputs selected at upper nodes of the regression tree. In addition, when an input is selected several times in different nodes of a regression tree, it implies that the input is of importance.

Figure 7-6 presents an example of CART results. The example comes from the case study scenario used in the slaughter module of the *E. coli* model (Frey *et. al.,* 2003). As shown in Figure 7-7, four inputs are selected in the regression tree, including chilling effect, number of organisms, washing effect, and contaminated surface. Because no constraint was placed on the number of nodes in the regression tree, the mean responses provided in the tree account for all the variability in the output that could be captured by partitioning the dataset. The classification rule for the highest mean combo bin contamination level corresponds with cases in which the

Table 7-6. Reduction in Deviance Associated with Selected Inputs in the Regression Tree Generated in the One-Dimensional Analysis of Variability and Uncertainty in the Slaughter Module

Inputs	Reduction in the Total Deviance Corresponding to Each Input							
	1st Level	2nd Level	3rd Level	4th Level	5th Level	Sum	Percent Contribution	Rank
Chiller Growth	5,600		1,770 (1)			7370	31.7	2
Number of Organisms		14,044(2)		53		14,097	60.6	1
Washing Effect				321		321	1.4	4
Contaminated Surface					1,464	1,464	6.3	3

(1) From this amount, 227 and 1,543 are associated to selection of chilling effect in the left and right branches of the tree, respectively.
(2) From this amount, 464 and 13,580 are associated to selection of number of organisms in the left and right branches of the tree, respectively

initial number of organisms on carcasses is greater than 128 organisms and growth in the chiller is higher than 2.2 logs. The mean response for these cases is 3641 *E. coli* organisms or approximately 3.6 logs of contamination.

Based on the regression tree, the combo bin contamination is most sensitive to both growth in the chiller and the initial number of organisms. The chilling growth is placed at the first node of the tree. However, although the number of organisms was not selected until the second level of the tree (in the right-most branch), this input discriminates the mean response of 3,641 from other leaves with mean responses of 130 to 528. Thus, CART suggests that the initial number of organisms accounts for a wide range of variation in the response. Therefore, there appears to be an important interaction between the chiller growth and the initial number of organisms. For low values of the chiller growth the mean response varies from 2.3 to 1358, depending on more refined ranges of chilling effect, number of organisms, and washing efficiency. For high values of the chiller growth, the initial number of organisms is the most important input. Therefore, it implies in this case that the chiller growth and the initial number of organisms are of comparable importance, and the other inputs selected in the tree are of minor importance.

Frey *et. al.,* (2003) evaluated an alternative sensitivity index for CART using the contribution of an input to the reduction of total dataset deviance. This sensitivity index is not automatically produced by the software package and was calculated separately. As an example,

Table 7-6 summarizes the contribution of the selected inputs to the reduction of the total deviance for the regression tree shown in Figure 7-6. There are five levels in the regression tree. Except for the first level, there are multiple branches at a given level. Therefore, an input may appear several times under different branches of a given level. Each such appearance is denoted with a numerical entry in Table 7-6.

The total deviance of the dataset for the one-dimensional analysis of variability and uncertainty is 44,280. Approximately 53 percent of the total deviance is reduced by classifying the dataset using the regression tree. For each input, the total reduction in the deviance is based upon the cumulative effects of repeated splits in the regression tree. The rankings in Table 7-6 indicate that the number of organisms is the most important input, although this input is not selected at the first splitting node. Chiller growth is selected at the first splitting node and ranks 2nd. The largest individual reduction in the total deviance is associated with the selection of initial number of organisms at the second level of the tree.

7.4 Graphical Sensitivity Analysis Methods

The section discusses the presentation and interpretation of two graphical sensitivity analysis methods: scatter plots and conditional sensitivity analysis. Details regarding the two methods and procedures for application of them to food safety risk assessment models are given in Frey et al. (2003) and Section 6.3.1, respectively.

Graphical methods can be presented in the form of two or three dimensional graphs depicting variation of the output versus the variation of one or two inputs. When presenting the results of graphical methods, it is crucial to select an appropriate number of data points. A small number of points may not reveal the underlying functional pattern, such as a threshold or non-linearity. In contrast, a dense graph showing thousands of data points may be difficult to interpret. A model may behave differently in different ranges of variation of an input. Hence, an analyst may choose to change the scale of the axes to present the results for a desired range of input or output values.

A few examples are provided here from case studies performed by Frey *et al* (2003). The examples provided here demonstrate the capabilities of the graphical sensitivity analysis methods in identifying specific model characteristics, such as non-linearity, thresholds, and interaction.

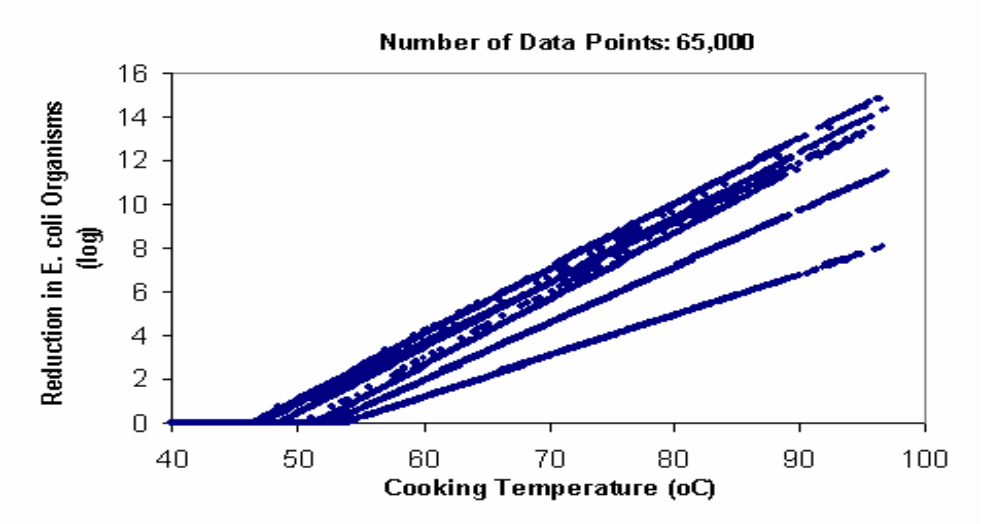

Figure 7-7. Scatter Plot for the Log Reduction in the Number of E. coli Organisms versus the Cooking Temperature at Home.

Figure 7-7 presents an example using the scatter plot method. The data shown in the figure are from a Monte Carlo simulation of the *E. coli* model. Figure 7-7 indicates a linear relationship between the log reduction in the number of *E. coli* organisms due to cooking and cooking temperature. Each line in the scatter plot represents a specific precooking treatment. That the lines presented in the scatter plot are not parallel indicates that there is an interaction between cooking temperature and the precooking treatment. Because of the interaction, the response of the model differs for low and high cooking temperature depending on precooking treatment. There is also threshold in the model response to cooking temperature. Cooking temperatures between 47°C and 53°C, depending on the precooking treatment, have no effect on the reduction in the number of *E. coli* organisms.

Figure 7-8 shows another example using the scatter plot. This figure indicates that there is an apparent threshold in the response of the model to the grinder contamination. When the grinder contamination is less than an approximate value of –2.5 logs, the grinder contamination has negligible effect on the contamination of the ground beef servings since less than one organism is predicted. In contrast, when the contamination in the grinder loads increases above the threshold value of –2.5 logs, there is a nonlinear relationship between the serving contamination and the grinder contamination, and ground beef servings become contaminated with more than one *E. coli* organism.

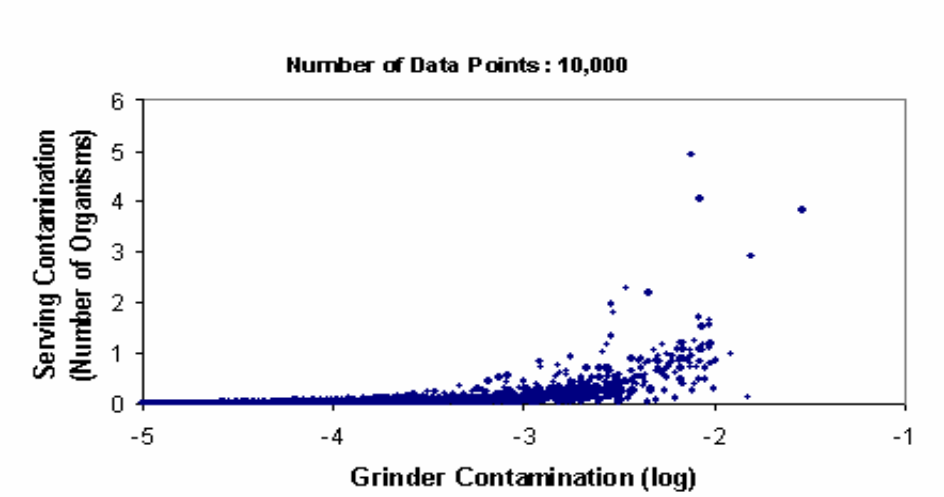

Figure 7-8. Scatter Plot for the Serving Contamination Versus the Grinder Contamination in Summer.

Figure 7-9 presents an example of a conditional sensitivity plot for the mean growth in the number of *E. coli* organisms versus storage time at stage 1. This figure not only reveals the way the model responds to variation of storage time conditioned on point estimate values of other inputs, but also identifies conditional thresholds. When other inputs are held at their mean values, there is no growth unless the storage time is greater than approximately 68 hrs. In this case, when the ground beef servings are stored for more than 68 hrs, there is a nonlinear response to the increase of the storage time. The approximate value of 68 hrs can be considered as a threshold in the model response to the variation of the storage time at stage 1, when other inputs are conditioned at their mean values. The threshold when other inputs are conditioned at their minimum values is approximately 86 hrs, which indicates that there is an interaction between the storage time and other inputs. If other inputs are held at their maximum values the threshold is approximately 4.5 hrs, and the maximum population density is reached after approximately 31 hrs. In contrast, when other inputs are held at their minimum or mean values, the maximum population density is not achieved even within 250 hours.

7.5 Summary

The presentation and interpretation of sensitivity analysis results are important to helping an analyst or a decision-maker better understand the insights from the results. This chapter first introduced the general principles for presenting and interpreting sensitivity analysis results. For each selected method in Chapter 5, an example is provided to help an analyst better understand

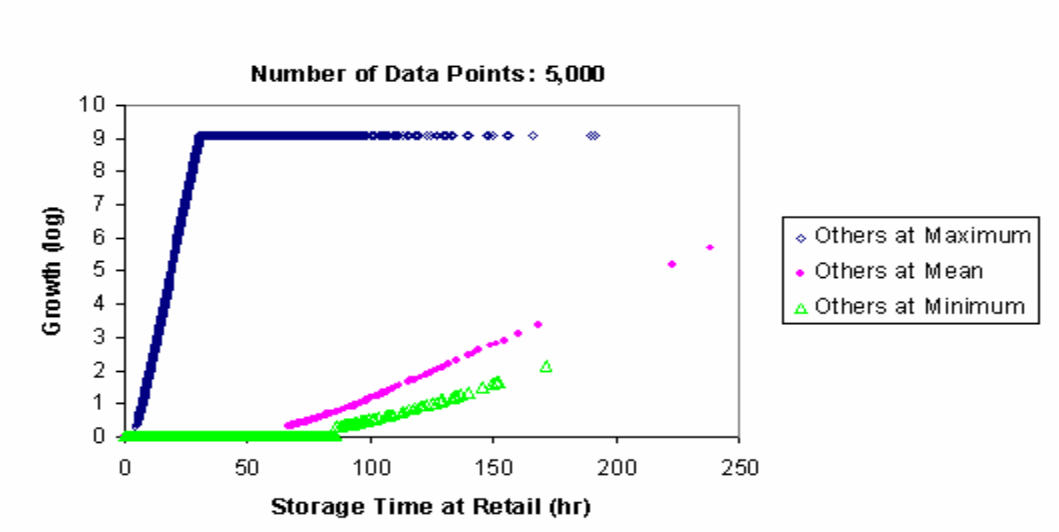

Figure 7-9. Conditional Sensitivity Analysis of Growth in the Number of *E. coli* Organisms to Storage Time at Retail.

the presentation and interpretation of the results. For any method, a key point in presenting and interpreting sensitivity analysis results is the necessity to clearly identify the target audience.

APPENDIX *A*. DISCUSSION OF ADDITIONAL STATISTICAL SENSITIVITY ANALYSIS METHODS

This appendix provides discussion of four statistical sensitivity analysis methods that are not commonly applied to food safety risk assessment models but that are potentially useful. These methods include: (1) Fourier Amplitude Sensitivity Test (FAST); (2) response surface method (RSM); (3) Mutual Information Index (MII); and (4) Sobol's method. These methods are briefly explained in Sections A.1 to A.4, respectively. For each method key features in application of the procedure and interpretation of results are discussed. Patil and Frey (2003) performed some case studies with FAST, RSM, and MII applied to the FDA *Vibrio parahaemolytics* model (CFSAN, 2001).

A.1 Fourier Amplitude Sensitivity Test (FAST)

FAST can be used for both uncertainty and sensitivity analysis (Cukier *et al.*, 1973, 1975, and 1978). FAST can identify the contribution of individual inputs to the expected value of the output variance (Cukier *et al.*, 1973). FAST does not assume a specific functional relationship such as linearity or monotonocity in the model structure, and hence works for monotonic and non-monotonic models (Saltelli *et al*., 2000). The effect of only one input or the effect of all inputs varying together can be assessed by FAST.

A.1.1 Description

FAST is a pattern search method that selects points in the input domain, and it is known to be faster than the Monte Carlo method (McRae *et al.*, 1982). The classical FAST method is not efficient in addressing higher-order interaction terms (Saltelli and Bolado, 1998). However, the extended FAST method developed by Saltelli *et al*. (1999) can address higher order interactions between the inputs.

FAST is used to estimate the ratio of the contribution of each input to the output variance with respect to the total variance of the output as the first order sensitivity index. This index can be used to rank the inputs (Saltelli *et al.*, 2000). Because FAST can allow arbitrarily large variations in input parameters, the effect of extreme events can be analyzed (*e.g.*, Lu and Mohanty, 2001; Helton, 1993). The evaluation of sensitivity estimates can be carried out independently for each factor using just *a single set of* simulations (Saltelli *et al*., 2000).

As a drawback in application of FAST, it suffers from computational complexity for a large number of inputs (Saltelli and Bolado, 1998). The classical FAST method is good only for

models with no important or significant interactions among inputs (Saltelli and Bolado, 1998). The reliability of the FAST method can be poor for discrete inputs (Saltelli *et al*., 2000). Current software tools for FAST such as SIMLAB, are not easily applied with existing food safety process risk models (Patil and Frey, 2003)

A.1.2 Application Procedure

FAST has been applied in fields such as performance assessment of waste disposal systems (*e.g*., Lu and Mohanty, 2001; Helton, 1993), atmospheric modeling (*e.g.*, Rodriguez-Camino and Avissar, 1998; Collins and Avissar, 1994; Liu and Avissar, 1996), and ground water modeling (Fontaine *et al*., 1992).

The main idea behind the FAST method is to use the properties of Fourier series to approximate the variance in output values and apportion the output variance to variance of the model inputs (Cukier *et al.,* 1973). Application of FAST involves defining a set of transformation functions and angular frequencies for model inputs. FAST uses the defined transformation function of each input for sampling during a probabilistic simulation of a model. Because FAST does not use the same probabilistic sampling technique as the one typically used in a food safety model (e.g., Monte Carlo sampling), application of FAST to a model will not be post hoc and the process should be integrated within the model. Some statistical software packages are available for application of FAST to a model. For example, SIMLAB has the ability to perform FAST (SIMLAB, 2000). Patil and Frey (2003) used SIMLAB for application of FAST to case scenarios with the FDA *Vibrio* model. Key steps in application of FAST for performing sensitivity analysis of a model with SIMLAB are:

- Encode the model in SIMLAB or encode the model in EXCEL and prepare a connection between SIMLAB and EXCEL using macros.
- Define distributions of model inputs in SIMLAB
- Generate sample values for each input using transformation functions in SIMLAB
- Perform a probabilistic simulation of the model using generated values for inputs from SIMLAB
- Estimate sensitivity indices using SIMLAB

A.1.3 Interpretation of the Results

FAST presents sensitivity in terms of the contribution of each input to the total output variance. The percentage contribution of each input to the total output variance can be estimated by normalizing the FAST indices for each input. FAST can provide: (1) first-order indices; (2) higher-order indices; and (3) total indices. Model inputs can be ranked using the relative magnitude of sensitivity indices.

For linear models the sum of the first-order indices of all inputs should equal to one. If the estimated sum does not equal one, it implies that interactions between inputs are statistically significant and contribute to the output variance. The total index for an input includes the first-order sensitivity index of the input and sum of the indices representing the interactions of the input with the other model inputs. The difference between the total sensitivity index and the first-order sensitivity index indicates the importance of the interaction effect. If the total index for an input is almost zero or a very low value, then this input can be fixed at a constant value without significantly impacting the results of the probabilistic simulation. Hence, the total indices can be used for simplifying the model.

A.2 Response Surface Method

The objective of the RSM is to develop a simplified version of the original model so that it is possible to retain the key characteristics of the model and to shorten the amount of time required to predict the output for a given set of inputs. RSM is typically applied to large models so that statistical methods that require multiple model evaluations can be applied. RSM is often used as a step prior to application of techniques that require many model evaluations, such as Monte Carlo simulation.

A.2.1 Description

A Response Surface (RS) can be linear or nonlinear and is typically classified as first-order or second-order methods (Myers and Montgomery, 1995). For nonlinear response surfaces interactions terms between inputs are considered. The number of inputs included in a RS and the type of RS structure required affect the amount of time and effort needed to develop a RS. It is often beneficial to limit the inputs that are included in the RS to those that are identified as most important using a screening sensitivity analysis method, such as NRSA.

A typical approach to RS development is to use a least-squares regression method to fit a standardized first or second order equation to the dataset including the output values from a

model and sampled values from probability distributions of model inputs. The precision and accuracy of the RS can be evaluated by comparing the prediction of the RS with those of the original model for the same values of the model input. The key assumption of least-squares regression, which is the normality of residuals, should be satisfied; otherwise, other techniques such as rank-based or nonparametric approaches should be considered (Khuri and Cornell, 1987; Vidmar and McKean, 1996).

Because the RS is calibrated to data generated from the original model, the valid domain of applicability of the RS model will be limited to the range of values used to generate the calibration dataset. Most RS studies are based on a fewer inputs than the original model. Therefore, the effect of all original inputs on the sensitivities cannot be evaluated in RSM.

A.2.2 Application Procedure

The application procedure for RSM starts with a decision regarding the inputs that should be included in the RS. If there are a large number of inputs, the RSM can be very complex. Hence, a practitioner can use the results from screening analysis or other sensitivity analysis methods to select a set of important inputs for the RS.

The next step after choosing important inputs included in RS is to verify the RS with the original data. The F test can be used to determine whether or not the RS is adequately accurate by comparing the variance of the output values generated by the RS and the variance of the actual values of the model output used to develop the RS.

Once a response surface is developed, the sensitivity of the model output to one or more of the selected inputs can be determined by: (1) inspection of the functional form of the response surface; (2) statistical analysis if regression analysis was used to develop the response surface; or (3) application of other sensitivity analysis methods to the response surface. The response surface can be thought of as a "model of a model" with an advantage of being simpler and faster to execute than the original model.

A.2.3 Interpretation of the Results

The estimated coefficients of the terms included in the RSM can be used for sensitivity analysis. The first order effect of inputs included in a RS can be ranked considering the relative magnitude of estimated RS coefficients as long as all inputs have been standardized. Ranking should be provided for statistically significant inputs that have P values less than 0.05 for the estimated coefficients.

A.3 Mutual Information Index

The objective of the Mutual Information Index (MII) sensitivity analysis method is to produce a measure of the information about the output that is provided by a particular input. The sensitivity measure is calculated based upon conditional probabilistic analysis. The magnitude of the measure can be compared for different inputs to determine which inputs provide the most information with respect to the output. MII is a computationally intensive method that takes into account the joint effects of variation in all inputs with respect to the output. MII is typically used for models with dichotomous outputs; but it can also be used for outputs that are continuous (Critchfield and Willard, 1986).

A.3.1 Description

The mutual information is a more direct measure of the probabilistic relationship of two random variables than other measures such as correlation coefficients (Jelinek, 1970). For example, the correlation coefficient of two random variables examines the degree of linear relationship between them. Two uncorrelated variables may not be independent; however, two variables with zero mutual information are statistically independent. In addition, the results from MII can be graphically presented. Calculation of the MII requires iterative application of Monte Carlo techniques that may lead to computational complexity, and thus make practical application difficult (Merz *et al.*, 1992). Critchfield and Willard (1986a) have suggested an approach using symbolic algebra, which is reported to be less computationally intensive. Because of the simplifying approximations that may be used in MII, the robustness of ranking based on the sensitivity measure is difficult to evaluate.

The mutual information between two random variables is the amount of information about a variable that is provided by the other variable (Jelinek, 1970). The average MII for each input (I_{XY}) is calculated based on the PDF of the input and on the overall and conditional confidence in the output. The amount of information about a variable that is provided by the variable itself is measured in terms of the "average self-information" (I_{YY}) of that variable. For the purpose of sensitivity analysis, a normalized measure of the MII (S_{XY}) is used which is the ratio of I_{XY} and I_{YY} (Jelinek, 1970).

A.3.2 Application Procedure

Application of MII involves three general steps (Critchfield and Willards, 1986): (1) generating an overall confidence measure of the output value; (2) obtaining a conditional confidence measure for a given value of an input; and (3) calculation of sensitivity indices.

In a model with dichotomous output, the probability of each output value is interpreted as the overall confidence for each value. As the first step to estimate the overall confidence on the output, all inputs are varied according to their respective distributions and the output distribution will be generated. The overall confidence can be estimated as the percentage of times the probability of the output is less than or equal to a defined value. The overall confidence on the output is estimated from its CDF.

In a conditional confidence analysis, an input of interest is held constant at one value and other inputs are varied according to their respective PDFs to generate the PDF for the output. Monte Carlo simulation is used in this computation.

A.3.3 Interpretation of the Results

Sensitivity of the inputs can be evaluated based on the relative magnitude of I_{XY} and S_{XY} values estimated for each input. S_{XY} indicates how much an input is important in providing the "statistical information" about the output as compared to the "statistical information" that output can provide about itself. Thus, S_{XY} may be a better measure to use when the relative importance of the inputs needs to be evaluated.

A.4 Sobol's Method

Sobol's methods (Sobol, 1990, 1993; Saltelli *et al*., 2000) are variance-based "global sensitivity analysis" methods based upon "Total Sensitivity Indices" (TSI) that take into account interaction effects. The TSI of an input is defined as the sum of all the sensitivity indices involving that input. The TSI includes both the main effect as well as interaction effects (Sobol 1990; Homma and Saltelli, 1996). For example, if there are three inputs A, B and C, the TSI of input A is given by S(A) + S(AB) + S(ABC), where S(x) is the sensitivity index of x. S(A) refers to the main effect of A. S(AB) refers to the interaction effect between A and B. S(ABC) refers to the interaction effect between A, B, and C.

A.4.1 Description

The use of Sobol's indices in the field of sensitivity analysis is new and there are few publications on the application of Sobol's indices as global sensitivity methods. Effort has been

made to reduce the computational complexity associated with calculation of Sobol's indices. Saltelli (2002a) discusses how to make the best use of model evaluations when calculating Sobol's sensitivity indices.

Sobol's method can cope with both nonlinear and non-monotonic models, and provide a truly quantitative ranking of inputs and not just a relative qualitative measure (Chan *et al*., 2000). The types of influence of an input that are captured by Sobol's method include additive, non-linear or with interactions. Furthermore, Sobol's method can be smoothly applied to categorical variables without re-scaling. Sobol (1990) and Saltelli (2002b) describe such an implementation.

Sobol's method, in general, is computationally expensive (Pastres et. al., 1999). It is difficult to apply Sobol's method to models with a large number of inputs and complex model structure such as modularity. There is no readily available software that facilitates application of Sobol's method.

A.4.2 Application Procedure

Given a model in the form of $y=f(X_1,X_2,\ldots,X_k)$, where the inputs X_i's are uncorrelated, y can be defined as the realization of a probabilistic process obtained by sampling each of the X_i from its respective probability distributions. Sobol's method defines sensitivity indices based on the decomposition of the output variance into terms due to either single input effect (first-order indices), or joint effects of more than one input.

Saltelli *et al.* (1993) proposed a Monte Carlo procedure for the estimation of the expected value of the output variance due to variation of the selected input using two sample matrices from the original probability distributions of inputs. These matrices are referred to as "sample" and "re-sample" matrices. The proposed procedure by Saltelli decreases the computational time associated with estimation of full set of first-order sensitivity indices (Saltelli, 2002). Furthermore, in order to estimate the confidence intervals on calculated sensitivity indices, bootstrap technique should be used to estimate the 95 percent confidence intervals for the firs-order and the total sensitivity indices (Archer and Saltelli, 1997)

A.4.3 Interpretation of the Results

Inputs can be prioritized using the relative magnitude of Sobol's sensitivity indices. For a completely linear model, the sum of the first-order sensitivity indices for the model inputs should equal to one. When interaction effects between model inputs are significant, the summation of the first-order sensitivity indices will be less than one. For models that are highly non-linear

only, small portion of the output variance would be attributed to the main effect of inputs. Thus, when first-order sensitivity indices are substantially low, total sensitivity indices should be used for prioritizing model inputs. Estimated confidence intervals can be used to evaluate the ambiguity in rankings based on sensitivity indices. Inputs for which sensitivity indices have overlapping confidence intervals may be interpreted to have the same relative importance on the variance of the output.

APPENDIX *B.* GLOSSARY

Add-In Software Packages "Add-in" is a term used, especially by Microsoft, for a software utility or other program that can be added to a primary program. Examples of add-in software packages are Crystal Ball™ and @Risk.

Aggregation Refers to situations in which multiple numerical values are combined into one numerical value, such as sum or mean value. In some cases, aggregation reflects the underlying process.

Ambiguity of Ranking Degree of uncertainty in estimated rankings of importance of model inputs from a sensitivity analysis method because of uncertainty in the sensitivity index or lack of significant differences in the sensitivity index of the inputs being jointly evaluated.

ANOVA Analysis of Variance, a statistically-based method that can be used for sensitivity analysis.

Binning A summarization technique in which values of an internal input or output of interest are binned into predefined intervals. Binning leads to loss of one-to-one correspondence between an output of a module and its exogenous predecessor inputs. If binning exists in a model, it can affect the scope of sensitivity analysis.

Bootstrap simulation A numerical method for estimating sampling distributions and confidence intervals for statistics based on an assumption that the "true" population distribution is equal to the empirical sample distribution.

Branch Term used in CART to refer to the conditions on the input variables that determine which output data go into particular data strata in a regression tree.

CART Classification and Regression Tree, a statistically-based method that can be used for sensitivity analysis. Also known as Hierarchical-Based Regression Tree (HBTR).

Categorical Input Also referred to as a qualitative input. This is a specific type of inputs with values that are mutually exclusive. The categories are measured in nominal scale and cannot be ordered from "highest" to "lowest" in any meaningful way. Rather, the categories define different groups. Examples include gender, season, and consumption type.

CCP Critical Control Point (CCP). A CCP is defined as a point, step, or procedure at which control can be applied and food safety hazard can be prevented, eliminated, or reduced to an acceptable level.

CDF The cumulative distribution function for a variable indicates the probability that the variable is less than or equal to any particular value.

CCDF Complementary cumulative distribution function for a variable indicates the probability that the variable is greater than any particular value.

Classification Rule With respect to CART, a result obtained from CART that enables stratification of sample values of the output to a specific situation determined by specific cut-points for specific inputs. For example, a classification rule could provide insight into conditions that lead to the largest mean exposure.

Coefficient of Determination (R^2) A statistic that is widely used as a goodness-of-fit measure for regression analysis and ANOVA. R^2 represents the fraction of variation in the output that can be explained by the fitted regression or ANOVA model.

Conditional Analysis Conditional analysis includes the "what-if" scenario analysis of a model and also identification of factors contributing to high exposures and risks.

Confidence Interval The computed interval with a given probability (i.e., 95%) that the true value of the statistic, such as a mean, proportion, or rate, is contained within the interval.

Contrast A comparison between two levels (or two sets of levels) of an independent variable in Analysis of Variance (ANOVA).

Correlation Coefficient Sample correlation coefficients provide a measure of linear association between two variables. Rank correlation coefficients provide a measure of monotonic association between two variables.

Critical limit A term used in HACCP to refer to a criterion or range that must be met for each preventive measure associated with a CCP, such as minimum internal temperature, product dimension, and cooking time.

CSA Conditional Sensitivity Analysis is a graphical method, in which the sensitivity of the model to a small number of inputs is evaluated while other inputs are held at fixed values.

Cut-Point With respect to CART, a value of an input used as the basis for creating branches.

Discrete Input An input variable that can only assume certain integer values. An ordinal categorical input variable could also be called discrete variable.

DSA Differential Sensitivity Analysis, a mathematical local sensitivity analysis method based upon small perturbations of inputs at a specific point in the model space.

F Value The F value is the ratio of two sums of squares (i.e. estimates of a population variance, based on the information in two or more random samples). When employed in ANOVA, the obtained F value provides a test for the statistical significance of the observed differences among the means of two or more random samples.

Factor An independent variable or input in ANOVA.

Factorial Design An experimental design in which each level of a variable is paired with each level of every other variables.

Food Safety Process Risk Model A quantitative model evaluating the risk to humans exposed to specific food borne hazard on a population basis. Food safety process risk models aim to simulate the process of bringing food from the farm to the table.

Global Sensitivity Analysis Global sensitivity analysis apportions the uncertainty in the output to the uncertainty in the input factors. Global methods are applicable to situations in which model inputs are varied simultaneously over large ranges of values, typically based upon probability distributions assigned to each input.

Graphical Sensitivity Analysis Methods A category of sensitivity analysis methods that feature the use of graphical techniques to present how a model output responds to changes in a model input. Typical graphical techniques include scatter plots or conditional sensitivity analysis.

HACCP Hazard Analysis and Critical Control Points (HACCP). HACCP is a management system in which food safety is considered through analysis and control of biological, chemical and physical hazards including raw materials, handling, production, distribution and consumption of the final product.

Important Inputs. An input to which the model output is highly sensitive or that leads to an outcome of most interest to a decision maker.

Interaction A case in which the effect of an input to the model response depends on the value of another input. For example, interactions can be represented in regression and ANOVA in the form of cross-product terms, such as $X_i \times X_j$, where X_i and X_j are inputs to the model.

Intermediate Node With respect to CART, an intermediate node refers to a node at which the data can be successively split in a regression tree.

Internal input A variable that is an output from a predecessor module and an input to a successor module for modules that are in series.

Leaf Node With respect to CART, leaf node refers to the node at which the penultimate data was split in a regression tree. The leaf node is also referred to as a terminal node.

Level A concept used in ANOVA to refer to specific range of values for a factor.

Local Sensitivity Analysis Local sensitivity analysis focuses on the impact of changes in input values with respect to a specific point in the input domain. NRSA and DSA are examples of local sensitivity analysis methods.

Mathematical Sensitivity Analysis Method A category of sensitivity analysis methods that are typically applied to a deterministic mathematical model. Examples include NRSA and DSA.

Measure of Sensitivity An index by which sensitivity of model output to inputs can be prioritized or by which insight regarding the sensitivity of the output to inputs can be obtained. Examples include standardized regression coefficients in regression analysis and F values in ANOVA. Also referred to as sensitivity index.

Model Independent A sensitivity analysis method is said to be model independent if it does not require any specific assumption regarding the functional form of the model to which it is applied.

Model Validation Model validation is the comparison of model results to independent observations from the system which is being modeled.

Model Verification A process of ensuring that the mathematical structure of a model, its computer implementation, and its input assumptions are as intended.

Monotonic Relationship A relationship that occurs when there is an association between two variables that is either consistently increasing or consistently decreasing. For example, the degree of monotonic relationship is measured by the Spearman Correlation Coefficient.

Node With respect to CART, a node is an input variable at which data can be split in a regression tree.

Nominal Value A selected point value of a distribution, such as a minimum, mean, median, or maximum value.

Non-linearity A relationship between two variables in which the change in each variable is not simply proportional to the change in the other variable. Parabolic and exponential relationships are examples of non-linearity. Non-linearity may be diagnosed from bivariate scatter plots.

NRSA Nominal Range Sensitivity Analysis – a mathematical technique in which one input at a time is varied over a range from low to high, while other inputs are held constant at a nominal value.

P Value The probability of obtaining a more extreme value for a test statistic under the assumption that the null hypothesis is true. The P value is compared with the chosen significance level of the test and, if it is smaller, the result is said to be significant.

Path or Classification Rule With respect to CART, a set of conditions on the input variable from the root node leading to a leaf node in a regression tree.

Pearson Correlation Coefficient A type of correlation coefficient that evaluates the strength of linear association between paired random output samples of output and input values. It is also known as the sample correlation coefficient.

Probabilistic Analysis A numerical analysis in which frequency (or probability) distributions are assigned to represent variability (or uncertainty) in model inputs and random sampling techniques are used to generate model input samples that are then propagated through formula(s) of the model specification to obtain a sample from the distribution of an output variable(s) of interest.

"Push-Button" Techniques A term used to refer to sensitivity analysis techniques that are built-in features of commonly used software packages. For example, the ability to calculate sample correlation coefficients is a built-in feature of Crystal Ball™.

Qualitative Input See categorical input.

Quantitative Input An input that takes numerical values. Quantitative inputs can be continuous or discrete. Examples include income, age, temperature, and time.

Rank Correlation Coefficient See Spearman correlation coefficient.

Realization A realization refers to one model simulation based upon one randomly sampled value for each probabilistic input.

Refined Analysis An analysis that is intended to be more comprehensive and/or accurate than a *screening analysis*. A refined analysis typically requires greater resources of time, effort,

and computation than a screening analysis, but is expected to provide more detailed and robust insights.

Response Surface Method A sensitivity analysis method that develops a simplified version of the original model so that it is possible to retain the key characteristics of the model and to shorten the amount of time required to predict the output for a given set of inputs.

Root Node With respect to CART, the first node at which data are split in a regression tree.

RTE Ready-to-Eat food. A food that can be consumed safely without further heat treatment, including reheating. Examples include deli salads, bagged salads, and hot dogs.

Sample Correlation Coefficient See Pearson correlation coefficient.

Sampling Distribution A probability distribution for a statistic, typically estimated based upon random sampling error. A sampling distribution is the basis for estimation of confidence intervals for a statistic.

Saturation point A saturation pint is a value for a model input above which there is no change in the output, but below which there can be substantial variation in the output. For example, the maximum population density.

Scatter Plot A figure in which individual data points are plotted in two-dimensional space.

Scenario A set of facts, assumptions, and inferences about a problem of interest that helps an analyst in evaluating, estimating, or quantifying the problem of interest.

Scenario Uncertainty Uncertainty reflecting that a defined scenario may fail to consider all the factors and conditions contributing to the variation of the output. In the case of the food safety process risk models, scenario uncertainty is typically associated with missing or incomplete information needed to fully define exposure and dose.

Screening Sensitivity Analysis Can be used to identify subset of inputs that controls most of the output variability with low computational effort especially in models that are computationally extensive and have a large number of inputs. Screening methods are typically less accurate than refined methods and may be used in tiered approach.

Sensitivity Analysis The assessment of the impact of changes in input values or assumptions on generated output values.

Sensitivity Index A quantitative measure of the degree of association between changes in the values of a model output relative to changes in the values of a model input

Simulation A series of calculations that attempt to predict a value or outcome. Simulation is useful for exploring outcomes that are not observable under the conditions of interest, or an outcome whose occurrence is potentially dangerous.

Spatial Dimension of a Scenario The spatial considerations for a scenario including the relevant geographic area. Depending upon the level of detail of the model, it may be necessary to incorporate multiple scales of geographic information.

Spearman Correlation Coefficient A measure of the strength of the monotonic relationship between two random variables. Thus, it is applicable to situations where there is a monotonic nonlinear relationship. It is also known as the rank correlation coefficient.

Statistical Sensitivity Analysis Methods A category of sensitivity analysis methods that are typically applied to probabilistic models for which random samples are generated for model inputs and corresponding samples are estimated for model outputs. Examples of statistical sensitivity analysis methods include correlation analysis, regression analysis, ANOVA, and CART.

Statistical Significance The satisfaction of a statistical criterion determined by a priori specification of an acceptable probability of wrongly rejecting a true hypothesis (i.e., usually 0.05 percent).

Threshold A value in an input domain below which a model output does not respond to changes in the input. An example is a specific storage temperature below which zero growth is estimated for an organism.

Temporal Dimension of a Scenario The temporal dimension of a scenario typically include: (1) the time for each major step in a process; (2) the activity patterns of individuals or agencies; (3) "seasonal" effects, whether at a short time scale (e.g., daily, weekly) or a longer scale (e.g., monthly, quarterly, annual); and (4) the time period associated with occurrence of illness as a result of one or more exposures.

Treatment With respect to ANOVA, a specific combination of levels for different factors. For example, in a factorial design with two factors that have n and m levels, respectively, $n \times m$ treatments can be specified.

Two-Dimensional Probabilistic Simulation A probabilistic simulation in which variability and uncertainty in model input are simulated separately.

Uncertainty Lack of knowledge about the "true" value of a quantity, lack of knowledge regarding which of several alternative models best describes a mechanism of interest, or lack of knowledge about which of several alternative probability density functions should represent a quantity of interest. Uncertainty is a property of an analyst.

Variability Heterogeneity of values over time, space, or different members of a population. Variability is a property of nature.

What-If Scenario Analysis A type of conditional analysis in which specific goals with respect to risk mitigation can be established. As an example of a "what-if" scenario, risk managers may be interested in decreasing the mean exposure level by a specific amount.

REFERENCES

Andersson, F.O., M. Aberg, and S.P. Jacobsson (2000), "Algorithmic Approaches for Studies of Variable Influence, Contribution and Selection in Neural Networks," *Chemo-metrics and Intelligent Laboratory Systems*, 51(1): 61-72.

Box, G.E.P, and G.C. Tiao (1973), *Bayesian Inference in Statistical Analysis,* Wiley-Interscience, New York.

Breiman (1984), "*Classification and Regression Trees*," Belmont, Calif.: Wadsworth International Group.

Brun, R., P. Reichert, and H.R. Kunsch (2001), "Practical Identifiability Analysis of Large Environmental Simulation Models," *Water Resources Research*, 37(4):1015-1030.

Buchanan, R.L., Smith J.L., and W. Long (2000), "Microbial Risk Assessment: Dose-Response and Risk Characterization," *International Journal of Food Microbiology*, 58(3):159-172.

Carlucci, A., F. Napilitano, A. Girolami, and E. Monteleone (1999), "Methodological Approach to Evaluate the Effects of Age at Slaughter and Storage Temperature and Time on Sensory Profile of lamb Meat," Meat Science, 52(4): 391-395.

Chan., K., S. Tarantola, A. Saltelli, and I. M. Sobol (2000), "Variance Based Methods," in *Sensitivity Analysis,* A. Saltelli, K. Chan, and M. Scott (Editors), John Wiley and Sons: New York. 167-197.

Cook, R.D (1994). *An Introduction to Regression Graphics.* John Wiley and Sons: New York.

Critchfield, G.C., and K.E. Willard (1986), "Probabilistic Analysis of Decision Trees Using Monte Carlo Simulation," *Medical Decision Making*, 6(1): 85-92.

CFSAN (2001), "Draft Risk Assessment on the Public Health Impact of *Vibrio parahaemolyticus* in Raw Molluscan Shellfish," Prepared by the *Vibrio parahaemolyticus* Risk Assessment Task Force (Center for Food Safety and Applied Nutrition) for Food and Drug Administration (FDA), USDA, Washington, DC.

CFSAN (2003), Quantitative Assessment of Relative Risk to Public Health from Foodborne *Listeria monocytogenes* Among Selected Categories of Ready-to-Eat Foods, Prepared by Center for Food Safety and Nutrition/Food and Drugs Administration/ U. S. Department of Agriculture, September 2003.

Cukier, R.I., C.M. Fortuin, K.E. Shuler, A.G. Petschek, and J.H. Schailby (1973), "Study of the Sensitivity of the Coupled Reaction Systems to Uncertainties in Rate Coefficients: I. Theory," *Journal of Chemical Physics*, 59(8): 3873-3878.

Cukier, R.I., J.H. Schailby, and K.E. Shuler (1975), "Study of the Sensitivity of Coupled Reaction Systems to Uncertainties in Rate Coefficients: III. Analysis of Approximations," *Journal of Chemical Physical*, 63(3): 1140-1149.

Cukier, R.I., H.B. Levine, and K.E. Shuler (1978), "Nonlinear Sensitivity Analysis of Multiparameter Model Systems," *Journal of Computational Physical*, 26(1): 1-42.

Cullen, A.C., and H.C. Frey (1999). *Probabilistic Techniques in Exposure Assessment.* Plenum Press: New York.

Devore, J.L., and R. Peck (1996). *Statistics: The Exploration and Analysis of Data*. 3rd Edition. Brooks/Cole Publishing Company: London, England.

DHHS (1998), *Federal Food Drug and Cosmetic Act*. Federal Register, U.S. Department of Health and Human Services. Washington, D.C. Vol. 63, No. 193.

Draper, N.R., and H. Smith (1981), *Applied Regression Analysis*. Second Edition. John Wiley and Sons: New York.

Edwards, A. L. (1976), "The Correlation Coefficient." Ch. 4 in *An Introduction to Linear Regression and Correlation.* San Francisco, CA: W. H. Freeman, pp. 33-46.

Efron, B. and R.J., Tabshirani (1993), *An Introduction to Bootstrap,* Chapman and Hall, New York.

ERS (2001), "Valuing the Health Benefits of Food Safety: A Proceedings." Report No. ERS MP 1570. Economic Research Service, United States Department of Agriculture. http://www.ers.usda.gov/publications/mp1570/, Viewed: May 10, 2001.

EPA (1997), " Guiding Principles for Monte Carlo Analysis," U.S. Environmental Protection Agency, EPA/630/R-97/001, March 1997.

Fontaine, D.D., P.L. Havens, G.E. Blau, and P.M. Tillotson (1992), "The Role of Sensitivity Analysis in Groundwater Risk Modeling for Pesticides," *Weed Technology*, 6(3): 716-724.

Frey, H.C. (2002), "Introduction to Special Section on Sensitivity Analysis and Summary of NCSU/USDA Workshop on Sensitivity Analysis," *Risk Analysis,* 22(3): 539-545.

Frey, H.C., A. Mokhtari, D. Tanwir (2003), "Evaluation of Selected Sensitivity Analysis Methods Based Upon Applications to Two Food Safety Process risk Models," Prepared by North Carolina State University for Office of Risk Assessment and Cost-Benefit Analysis, U.S. Department of Agriculture, Washington, DC , June 2003.

Frey, H.C., and D.S. Rhodes (1996), "Characterizing, Simulating, and Analyzing Variability and Uncertainty: An Illustration of Methods Using an Air Toxics Emissions Example," *Human and Ecological Risk Assessment: an International Journal*, 2(4): 762-797 (December).

Frey, H.C., and R. Patil (2002), "Identification and Review of Sensitivity Analysis Methods," *Risk Analysis,* 22(3): 553-577.

FSIS (2001), *Draft Risk Assessment of the Public Health Impact of Escherichia Coli O157:H7 in Ground Beef*, Prepared for the Food Safety and Inspection Service and United States Department of Agriculture, Washington, DC. http://www.fsis.usda.gov/OPPDE/rdad/FRPubs/00-023N/00-023NReport.pdf.

Galvao, L.S., M.A. Pizarro, and J.C.N. Epiphanio (2001), "Variations in Reflectance of Tropical Soils: Spectral-Chemical Composition Relationships from AVIRIS Data," *Remote Sensing of Environment*, 75(2): 245-255.

Geldermann, J., and O. Rentz (2001), "Integrated Technique Assessment with Imprecise Information as a Support for the Identification of Best Available Techniques (BAT)," *OR Spektrum*, 23(1): 137-157.

Gibbons, J. D. (1985). *Nonparametric Statistical Inference*. Statistics: textbooks and monographs. Marcel Dekker, Inc, New York and Basel, 2nd edition.

Giesbrecht, F. and Gumpertz, M. (1996), *Planning, Construction, and Statistical Analysis of Comparative Experiments*, Publication: John Wiley and Sons, Inc.

Helton and Davis (2001), "Illustration of Sampling-Based Methods for Uncertainty and Sensitivity Analysis*,"* Department of Mathematics, Arizona State University, Tempe.

Homma, T., and A. Saltelli (1996), "Importance Measures in Global Sensitivity Analysis of Nonlinear Models," *Reliability Engineering and System Safety*, 52(1): 1-17.

Hulebak, K.L, and W.Schlosser (2002), "Hazard Analysis and Critical Control Point (HACCP) History and Conceptual Overview," *Risk Analysis*, 22(3): 553-578.

Iman, R.L., M.J. Shortencarier, and J.D. Jhonson (1985). "A FORTRAN 77 Program and Users Guide for the Calculation of Partial Correlation and Standardized Regression Coefficients." Report No. SAND85-0044, Sandia National Laboratories, Albuquerque, NM.

Jelinek, F. (1970), *Probabilistic Information Theory,* McGraw Hill Book Company: New York.

Kendall, M. and A. Stuart (1979), *The Advanced Theory of Statistics*, 4th Ed. MacMillian Publishing Co.: New York.

Khuri, A.J., and J.A. Cornell (1987), *Response Surfaces.* Marcel Dekker, Inc.: New York.

Kleijnen, J.P., and J.C. Helton (1999), "Statistical Analysis of Scatter plots to Identify Important Factors in Large-Scale Simulations," Sandia National Laboratory, SAND98-2202, Albuquerque, New Mexico.

Lammerding, A.M. (1997), "An Overview of Microbial Food Safety Risk Assessment," *Journal of Food Protection*, 60(11): 1420-1425.

Lu, Y.C., and S. Mohanty (2001), "Sensitivity Analysis of a Complex, Proposed Geologic Waste Disposal System Using the Fourier Amplitude Sensitivity Test Method," *Reliability Engineering and System Safety*, 72(3): 275-291.

McCamley, F., and R.K. Rudel (1995), "Graphical Sensitivity Analysis for Generalized Stochastic Dominance," *Journal of Agricultural and Resource Economics*, 20(2): 403-403.

Mead, P.S., S. Laurence, D. Vance, L.F. McCaig, J.S. Bresee, C. Shapiro, P.M. Griffin, and R.V. Tauxe (1999), "Food-Related Illness and Death in the United States," *Emerging Infectious Diseases*, 5(5): 607-625.

Merz, J.F., M.J. Small, P.S. Fischbeck (1992), "Measuring Decision Sensitivity: A Combined Monte Carlo-Logistic Regression Approach," *Medical Decision Making,* 12(3): 189-196.

Morgan, M.G., and M. Henrion (1990). *Uncertainty: A Guide to Dealing With Uncertainty in Quantitative Risk and Policy Analysis.* Cambridge University Press: Cambridge, NY.

Myers, R.H., and D.C. Montgomery (1995), *Response Surface Methodology: Process and Product Optimization Using Designed Experiments.* Wiley and Sons Ltd.: New York.

Nauta, M.J. (2000), "Separation of Uncertainty and Variability in Quantitative Microbial Risk Assessment Models," *International Journal of Food Microbiology* 57: 9-18.

Nauta, M.J. (2001a), "A Modular Process Risk Model Structure for Quantitative Microbiological Risk Assessment and Its Application in an Exposure Assessment of *Bacillus Cereus* in a REPFED," RIVM Report 149106 007, National Institute for Public Health and the Environment, Bilthoven.

Nauta, M.J. (2001b), "Modeling Bacterial Growth in Quantitative Microbiological Risk Assessment: Is It Possible?," *International Journal of Food Microbiology*, 73(2-3): 297-304.

Neter, J., W. Wasserman , and M.H. Kutner, (1989), *Applied Linear Regression Models*, Second Edition. IRWIN: Homewood, IL.

Neter, J., M.H. Kutner, C.J. Nachtsheim, and W. Wasserman (1996), *Applied Linear Statistical Models*, Fourth Edition. McGraw-Hill: Chicago, IL.

Pastres, R., K. Chan, C. Solidoro and C. Dejak (1999), *"*Global sensitivity analysis of a shallow-water 3D Eutrophication model*," Computer Physics Communications,* 117(1-2): 62-74. Elsevier (Amsterdam).

Rose, J.B. (1993). "Waterborne Pathogens: Assessing Health Risks," *Health and Environment Digest*, 7(3): 1-3.

Saltelli, A., and R. Bolado (1998), "An Alternative Way to Compute Fourier Amplitude Sensitivity Test (FAST)," *Computational Statistics and Data Analysis*, 26(4): 445-460.

Saltelli, A., S. Tarantola, and K. Chan (1999), "A quantitative, Model Independent Method for Global Sensitivity Analysis of Model Output," *Technometrics*, 41(1): 39-56.

Saltelli A., K. Chan, and M. Scott, Eds. (2000), *Sensitivity Analysis*, Probability and Statistics Series. John Wiley & Sons: New York.

Saltelli, A. (2002a), "Sensitivity Analysis for Importance Assessment," *Risk Analysis*, 22(3): 1-12.

Saltelli, A. (2002b), "Making Best Use of Model Evaluations to Compute Sensitivity Indices," *Computer Physics Communication,* 145(2): 280-297.

Sen, A., and M. Srivastava (1990). *Regression Analysis: Theory, Methods, and Applications*. Springer-Verlag: New York.

Seward, S. (2000), "Application of HACCP in Food Service," *Irish Journal of Agricultural and Food Research*, 39(2): 221-227.

Siegel, S. and N.J. Castellan (1988). *Nonparametric Statistics for the Behavioural Sciences* (2nd edn), McGraw-Hill, New York.

SIMALB (2000), Software for Uncertainty and Sensitivity Analysis User Manual. POLIS-JRC/ISIS

Steel, R.G.D., J.H. Torrie, and D.A. Dickey (1997). *Principals and Procedures of Statistics; A Biometric Approach*. 3rd Edition. WCB McGraw-Hill: Boston, Massachusetts.

Sobol, I. M. (1990), "Sensitivity Estimates for Nonlinear Mathematical Models," *Matematicheskoe Modelirovanie,* 2(1): 112-118 (in Russian, translated as I. M. Sobol (1993), "Sensitivity Analysis for Nonlinear Mathematical Models," *Mathematical Modeling and Computational Experiment*, 1(4): 407-414).

Sobol, I.M. (1993), "Sensitivity Estimates for Nonlinear Mathematical Models," *Mathematical Modeling and Computation*, 1(4): 407-414.

Stiber, N.A., M. Pantazidou, and M.J. Small (1999), "Expert System Methodology for Evaluating Reductive Dechlorination at TCE Sites," *Environmental Science and Technology*, 33(17): 3012-3020.

Vidmar, T.J., and J.W. McKean (1996), " A Monte Carlo Study of Robust and least Squares Response Surface Methods," *Journal of Statistical Computation and Simulation,* 54(1): 1-18.

Vose, D.J. (2000), *Risk Analysis: A Quantitative Guide*, 2nd Edition, John Wiley & Sons Ltd. Chichester, England.

Washington, S., J. Wolf, and R. Guensler (1997), "A Binary Recursive Partitioning Method for Modeling Hot-Stabilized Emissions from Motor Vehicles," Prepared by School of Civil and Environmental Engineering, Georgia Institute of Technology for the 76th Annual Meeting of the Transportation Research Board, Atlanta, Georgia.

Zheng, J., H.C. Frey (2003), "Design and Development Report of the C++ Version of EPA/SHEDS/Pesticide Model," Prepared by Department of Civil, Construction and Environmental Engineering at North Carolina State University for the Office of Research and Development, U.S. Environmental Protection Agency, Research Triangle Park, NC, June 2003.