# Evaluation of Selected Sensitivity Analysis Methods Based Upon Applications to Two Food Safety Process Risk Models

Prepared by:

H. Christopher Frey
Amirhossein Mokhtari
Tanwir Danish
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

September 2003

## 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 reviewed and provided comments on a draft of this report:

John Bowers, Center U.S. Food and Drug Administration
Peter Cowen, North Carolina State University
David Dickey, 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, Office U.S. Department of Agriculture
Andrea Saltelli, European Commission Joint Research Centre
Wayne Schlosser, U.S. Department of Agriculture
Mark Walderhaug, U.S. Food and Drug Administration

We thank Wayne Schlosser of USDA and Clark Carrington of FDA for their assistance in installing and running the food safety risk assessment process models used as case studies in this work and for answering numerous questions regarding these models.

## Disclaimers

Any opinions, findings, conclusions, or recommendations expressed in this report 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

Concern for the safety of the food supply is motivated by recognition of the significant impact of microbial food borne diseases in terms of human suffering and economic costs to the society and industry (Lammerding, 1997). Mead et al. (1999) have reported that food borne disease results in 76 million human illnesses in the United States each year, including 325,000 hospitalizations and 5200 deaths. ERS (2001) estimated the cost of food borne disease to be $6.9 billion annually. The increase in international trade in food has increased the risk from cross-border transmission of infectious agents and underscores the need to use international risk assessment to estimate the risk that microbial pathogens pose to human health. The globalization and liberalization of world food trade, while offering many benefits and opportunities, also presents new risks. Because of the global nature of food production, manufacturing, and marketing, infectious agents can be disseminated from the original point of processing and packaging to locations thousands of miles away.

Food Safety regulatory agencies ensure the safety of the food supply based upon the Hazard Analysis Critical Control Points (HACCP) system (Hulebak and Schlosser, 2002; Seward, 2000). One step in the HACCP system is to determine critical control points (CCP) where risk management efforts can be focused. As an example, a critical control point could be selection of storage temperature or storage time of a food so as to prevent significant growth of microbial pathogens. Because of the complex and dynamic nature of the food processing, transportation, storage, distribution, and preparation system, identification of the critical control points in a farm-to-table pathway poses a substantial analytical challenge (Rose, 1993; Buchanan et al. 2000).

Sensitivity analysis of risk models can be used to identify the most significant exposure or risk factors to aid in developing priorities for risk mitigation. 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. Sensitivity Analysis can also be used to provide insight into the robustness of model results when making decisions (Cullen and Frey, 1999).

## 1.1 Objectives

The objectives of this project are to identify, review, and evaluate sensitivity analysis methods based upon case studies with two food safety risk assessment models developed by

USDA and FDA. This objective serves as an aid in identifying potential control points along the farm-to-table continuum, to inform decisions about food safety research and data acquisition priorities, and to contribute to the development of sound food safety regulations. This project is follow-on to previous work as described in Section 1.4. The main focus here is on the methodology of performing sensitivity analysis and regarding the key insights that such analysis affords.

The key questions that must be addressed in performing sensitivity analysis with food safety risk assessment models include the following (Frey, 2002):

- What are the key criteria for sensitivity analysis methods applied to food-safety risk assessment models?
- What sensitivity analysis methods are most promising for application to food-safety risk assessment models?
- What are the key needs for implementation and demonstration of such methods?

To address these questions, multiple sensitivity analysis methods were explored and applied to food safety process risk assessment models for *E. coli* O157:H7 and *Listeria monocytogenes*. This report presents application of several sensitivity analysis methods to these models. The analyses of this work are targeted to answer several key questions that address the overall project objectives. The answers to these questions are discussed in the conclusion chapter. For example, some specific questions include:

- Can simple sensitivity analysis methods such as nominal range sensitivity analysis provide robust insights in spite of their apparent limitations?
- Which methods can take care of qualitative and quantitative variables simultaneously?
- Which methods can identify and appropriately respond to thresholds?
- Which methods can specifically address high exposure/risk case scenarios?
- Which methods can give insights on interactions between explanatory variables?
- Which methods can identify or appropriately deal with non-linearity in response?
- How unambiguous is the relative importance of the model inputs based on the selected sensitivity index?
- How should sensitivity analysis be conducted in a two-dimensional probabilistic framework?

- Which sensitivity analysis methods can be easily automated to address the additional complexity introduced by two-dimensional Monte Carlo simulation of variability and uncertainty?

## 1.2 Food Safety Risk Assessment Modeling

Risk can be represented as a combination of the probability of occurrence and the impact of adverse effects caused by a hazard. Risk assessment is the process of identifying a hazard and qualitatively or quantitatively presenting the estimated risk of the hazard. Recently, risk assessment has been gaining support in governments worldwide as a mechanism for improving decision making and foretelling regulatory policy effects on public health. The U.S. government has made commitments to use risk assessment for many different types of decisions including food safety (WHO, 1995).

The traditional model of toxicological health risk assessment consists of four steps: (1) hazard identification, (2) hazard characterization, (3) exposure characterization, and (4) risk characterization (WHO, 1995). Each if these four steps are described briefly.

- Hazard Identification. Hazard identification involves listing biological, physical or chemical hazards of concern to human health that may be associated with the commodity/product situation in question, or conditions that alter the probability of significant human exposure to such disease agents. Scenarios may be very specific, describing food type, processing, potential contamination, storage, preparation methods, pH, water activity, temperature and other factors. The process of hazard identification may involve data searches, literature reviews, consultation with municipal, provincial, national or international organizations or university or industry experts including technical personnel involved directly with the product.
- Hazard Characterization. Hazard characterization is the qualitative and/or quantitative evaluation of the nature of the adverse health effects associated with biological, chemical, and physical agents that may be present in food. Hazard characterization may or may not include dose-response assessment.
- Exposure Characterization. Where possible, substantiated evidence may be used to build quantitative multiplicative models, to help estimate the probability of people experiencing the negative impact of a food borne health hazard. Stochastic and/or deterministic models may be used. Stochastic models mimic natural variability by

including a process of random selection within defined probability distributions. Deterministic models calculate overall probabilities based on a series of point estimates and do not include a process of random selection. The uncertainty of evidence is modeled by widening the distribution boundaries set in stochastic models, or by altering point estimates in sensitivity analysis of deterministic models.

- Risk Characterization. Risk is characterized by estimating in qualitative or quantitative terms, the probability of and the magnitude of the impact (or consequence) of the adverse effects of the disease for individuals and for a population. The risk is further characterized by noting the attendant uncertainty of the estimates, given the available data.

## 1.3 Need for Sensitivity Analysis

Sensitivity analysis is the assessment of the impact of changes in input values on model outputs (Cullen and Frey, 1999). In combination with uncertainty analysis, sensitivity analysis can include the study of how uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model inputs. Hence, sensitivity analysis is considered by some as prerequisite for model building in any setting, whether diagnostic or prognostic, and in any field where models are used. Quantitative sensitivity analysis is increasingly invoked for corroboration, quality assurance, and validation of model-based analysis (Saltelli, 2002).

Sensitivity analysis can be helpful in verification of a model. Verification is a process of checking that the model is correctly implemented. If a model responds in an unacceptable way to changes in one or more inputs, then trouble-shooting efforts can be focused to identify the source of the problem. Sensitivity analysis can be used to evaluate how robust risk estimates and management strategies are to model input assumptions and can aid in identifying data collection and research needs (Frey and Patil, 2002).

## 1.4 Summary of Previous Work at NCSU

On June 11-12, 2001, NC State University hosted a Workshop on Sensitivity Analysis, sponsored by the U.S. Department of Agriculture's Office of Risk Assessment and Cost Benefit Analysis (USDA/ORACBA). The workshop was part of a project whose objective was to transfer, apply, and adapt sensitivity analysis methods developed in other disciplines (e.g.

complex engineering systems) to food-safety risk assessment. The workshop proceedings have been published as a special section of the journal Risk Analysis.

A guest editorial in Risk Analysis describes the HACCP concept that underlies risk assessment and risk management pertaining to food safety (Hulebak and Schlosser, 2002). Because the workshop was comprised of participants with different disciplinary backgrounds, it was important to introduce everyone to a similar conceptual framework.In order to learn from different disciplines, and in preparation for the workshop, NCSU prepared a literature review regarding sensitivity analysis methods, including the strengths and limitations of selected methods that merit consideration for possible application to food safety risk assessment (Frey and Patil, 2002). The report presents a brief overview of the risk assessment framework pertaining to food safety risk assessment and then reviews key issues in food safety risk modeling, including the purpose of the model, complexity, verification, validation, extrapolation, and the role of sensitivity analysis. Sensitivity analysis methods are classified as mathematical, statistical, and graphical. Ten specific methods are reviewed, including nominal range sensitivity analysis, difference in log-odds ratio, break-even analysis, automatic differentiation, regression analysis, analysis of variance, response surface methods, Fourier amplitude sensitivity test, mutual information index, and scatter plots. For each method, a description, example, advantages, and disadvantages are addressed. The methods are compared with respect to applicability to different types of models, computational issues, ease and clarity in representation of results, and purpose of the sensitivity analysis. Some methods are model-free and global in nature, and may be better able to deal with non-linear models that contain thresholds and discrete inputs than can other methods. However, because each sensitivity analysis method is based upon different measures of sensitivity, two or more methods can in general produce dissimilar results. Therefore, as a practical matter, it is advisable to explore two or more techniques. Each method is good at extracting one or more features of the problem, and each feature corresponds to a different question put to the system. Setting up appropriately suited sensitivity analysis is discussed in Saltelli and Tarantola, 2002. Examples of questions that can be asked are:

- Which factor produces fractionally the greatest increment of the output?
- Which factors contribute the most to the variance of the output?
- Which factor is mostly responsible for producing realizations of the output beyond the $95^{th}$ percentile of the distribution of the output (or above a given threshold)?

While the answer to the first question is often the goal of sensitivity analysis, the second and third questions can be addressed by variance based methods and methods that aid in characterization of specific case scenarios (e.g., high end exposure cases), respectively.

Selected experts were invited to write and present "white papers" reviewing the application of sensitivity and/or uncertainty analysis to complex engineered and/or environmental systems. The purpose of these white papers was to: (1) summarize the development of sensitivity and uncertainty analysis of complex simulation methods in order to synthesize lessons learned in the field; (2) provide a state-of-the-art review and critique of selected applied methods and approaches; and (3) identify the most promising methods and approaches for application to large, complex food safety process risk models. Each of the five papers is briefly summarized here.

The first paper highlighted important criteria for sensitivity analysis methods (Saltelli, 2002). These included the need to properly specify a model output that is directly relevant to a decision, as well as identification of desirable properties in sensitivity analysis methods. The latter includes ability to cope with the scale of inputs and the shape of distributions assigned to inputs; global methods that can deal with the simultaneous effects of variation in multiple inputs; model independent methods that work regardless of the functional form of the model; and an ability to group inputs as if they were a single factor. A distinction was made between prognostic (forecast) and diagnostic (estimation) models. Variance-based methods, such as variations of Sobol's method, are described and illustrated with an example using a prognostic model.

The second paper illustrates the use of Latin Hypercube sampling combined with statistical and regression techniques in an overall approach for first propagating probability distributions through a model and then analyzing the results to identify the most sensitive inputs (Helton and Davis, 2002). With 150 cited references, the paper also provides the reader with an introduction to a large supporting literature.

The third paper discusses the reliability of a model, which in the author's view is related to the testability of the model (Kohn, 2002). The author introduces sensitivity analysis techniques based upon system sensitivity theory, with applications to empirical models and to metabolic networks. Examples of the application of such methods to physiological modeling are reviewed, illustrating the dynamic nature of sensitivities. Sensitivity analysis was shown to

provide insight into the apportionment of the model response to various inputs in a manner that can be explained based upon understanding of the biological processes being modeled.

The fourth paper places risk analysis and sensitivity analysis more squarely in the context of government decision-making, including the process of formulating hypotheses and bounding of the risk analysis problem (Pate-Cornell, 2002). A probabilistic framework based upon Bayesian methods is described. This approach is motivated because "expert judgment is simply unavoidable" in most risk assessment problems. This paper places the need for and interpretation of sensitivity analyses in the context of the formulation of a risk problem, including the scenarios and the model, the source of information for developing model inputs, and the specific methods used to model the risk problem.

The fifth paper addresses the risk management implications of the trend from point-estimate risk analysis to analyses that explicitly address both variability and uncertainty (Thompson, 2002). Using two example case studies, one based upon ground fatalities attributable to airline crashes, and the other based upon the risks and benefits of airbags, Thompson illustrates the importance of explicitly accounting for variability in risks. With the growing role of probabilistic risk assessments pertaining to food safety, as reflected by recent examples for foodborne *Listeria monocytogenes*, Vibrio parahaemolyticus in raw molluscan shellfish, Campylobacter in chicken, *E. coli* O157:H7 in beef, and Salmonella Enteritidis in shell eggs and egg products, there will be a need for risk managers to take into account both variability and uncertainty when developing risk management strategies.

The group did not recommend specific sensitivity analysis methods for application to food safety risk assessment models. Instead, the group emphasized using the methods that can deal with model characteristics like interactions, nonlinearities, discontinuities, and discrete inputs. The group recommended use of two or more sensitivity analysis method to obtain insight into robustness of results. Overall, the workshop resulted in identification of key criteria for sensitivity analysis methods and recommendations for work needed, listed in Section 1.5, to further evaluate and specify appropriate sensitivity analysis approaches in the context of food-safety risk assessment.

## 1.5 Challenges for Sensitivity Analysis

The workshop addressed three key questions pertaining to the application of sensitivity analysis in food safety risk assessment. These three key questions are also the focus of this

report. The insights obtained from the workshop are briefly reviewed here for each question and are further discussed by Frey (2002).

*Question 1: What are the key criteria for sensitivity analysis methods applied to food-safety risk assessment models?*

The workshop participants agreed that a key criterion for sensitivity analysis and for the risk model and analysis in general, are that it must be relevant to a decision. This means that the model output of interest must be directly related to the decision. Using a highly stylized example, if a decision is informed by whether risk is above or below a threshold, then the model output should be a variable indicating the probability that the estimated risk is above or below the threshold. The sensitivity analysis should pertain to variation in inputs that cause a change in the value of the output that would lead to a different decision.

Technical requirements of a sensitivity analysis method are manifold and differ from one application to another, and from one decision application to another. The ideal sensitivity analysis method would be applicable to models that have the following characteristics that are typical of food safety risk models:

- Nonlinearities;
- Thresholds (e.g., below which there is no growth of a microbial pathogen);
- Discrete inputs (e.g., integer numbers of animals or herds, yes or no indicators of contamination);
- Incorporation of measurement error;
- Variation in the scale (units and range) and shape of the distributions of model inputs; and
- Temporal and spatial dimension, including dynamics, seasonality, or inter-annual variability.

An ideal sensitivity analysis method would be model independent i.e., functional form (e.g., monotonic). Specifically, the sensitivity analysis method should not require the introduction of any assumptions regarding the functional form of the risk model and, therefore, should be applicable to a wide range of different model formulations. The method should provide not just a rank ordering of key inputs, but also some quantitative measure of the sensitivity of each input so that it is possible to distinguish the most strongly sensitive inputs from those with weaker influence on the selected model output. For example, is the most

sensitive of the inputs substantially more important than the second ranked input, or do the top two inputs have approximately equal influence on the model output?

Another challenge regarding the application of sensitivity analysis methods to food-safety risk assessment models is the importance of distinguishing variability and uncertainty where appropriate. It should be noted that such a distinction could be useful but not essential in every case. Thus, it may or may not be necessary, in a particular assessment, to distinguish between variability and uncertainty when doing the sensitivity analysis. It is recommended that the key sources of uncertainty that are based on data analysis be distinguished from key sources of uncertainty that are based on expert judgment. For important uncertain inputs for which uncertainty was estimated based upon expert judgments, refinements can be made based upon additional expert elicitation or development of an appropriate data collection effort. Refinement of important estimates of uncertainty that were based upon data analysis would typically require collection of additional data.

*Question 2: What sensitivity analysis methods are most promising for application to food-safety risk assessment models?*

The workshop participants did not identify specific methods. Instead, the group emphasized the key criteria that were generated in response to the first question. For example, methods that can deal with interactions, nonlinearities, discontinuities, and discrete inputs would be preferred over methods that cannot. Methods that are global or generic, such as ANOVA, are likely to be more promising than other types of methods, although ANOVA also has some limitations. However, techniques are also needed that can identify not just the effect of variance in the inputs, but also shift in central tendency or position of the output associated with skewness of distributions assigned to inputs.

Before applying a sensitivity analysis method, it may help to reduce the computational burden by narrowing down the search space among the input parameters. For example, if adverse consequences do not occurred unless a storage temperature exceeds a threshold above which microbial growth becomes significant, it may not be necessary or important to analyze model behavior when the storage temperature is below the threshold. Thus the search space could be narrowed to cases where the storage temperature is above the threshold in order to reduce computational time.

The goal of this research work is to identify the sensitivity analysis methods that are most promising for application to food safety risk assessment models, based upon case studies with

two representative models in order to judge the practical applicability of the methods. If a method allows for particular features of interest such as non-linearity, discontinuities, and discrete inputs then it is preferred over others.

*Question 3: What are the key needs for implementation and demonstration of such methods?*

The workshop participants agreed that different methods of sensitivity analysis should be explored and applied to more than one food safety risk model. The methods should be tested at research institutes and efforts should be made to confirm and validate the results. The process of testing methods will help establish a track record for specific methods applied to food safety process risk models. A comparison of methods, taking into account real life constraints, should be part of the guideline. The guideline should outline a tiered approach to sensitivity analysis.

Because sensitivity analysis formulation is conditional on the assumption that the model formulation is acceptable, it is important to have a prior comfort with the plausibility of the model and to examine the sensitivity analysis results to determine if any of the model responses are inconsistent with plausible expectations regarding the relationship between the model output and model inputs.

## 1.6 Selection of Models for Case Studies

Based on the results of the workshop on sensitivity analysis held at NC State it was decided that multiple sensitivity analysis methods should be applied to two food safety risk assessment models, namely Escherichia Coli O157:H7 in ground beef and *Listeria monocytogenes* among selected categories of ready-to-eat foods. The former model was developed by the U.S. Department of Agriculture, while the latter was developed by the Food and Drug Administration (FSIS, 2001 and CFSAN, 2001). These models are referred to here as the "*E. coli*" and "*Listeria monocytogenes*" models, respectively. These models are further described in Chapter 3 and Chapter 12, respectively.

A key feature that distinguishes the two selected food risk assessment models is that there was no clear objective or risk management question posed in the *E. coli*:O157 study. However in the case of *Listeria monocytogenes* study the objective was to arrive at the relative risk ranking for various food groups considered and thus prioritize future risk reduction efforts among the food categories.

Both models are non-linear and have thresholds. The *Listeria monocytogenes* model has an upper limit on possible growth of *Listeria monocytogenes*. The *E. coli* model has threshold in

the growth estimation part. For example, there is no growth below a particular storage temperature. For the *Listeria monocytogenes* model, all inputs are continuous, while the *E. coli* model has both discrete and continuous inputs. For example, the storage temperature is a continuous input in both models, whereas the ground beef consumption type is defined as a discrete input in *E. coli* model. The two models do not have a spatial dimension in that they do not explicitly account for differences in geographic location. The *Listeria monocytogenes* model addresses the risk to three sub-populations, namely, neonatal, intermediate and elderly, while the *E. coli* model considers two different cattle categories that act as *E. coli* sources or carriers. The *E. coli* model also has a temporal dimension in that high and low prevalence seasons are considered separately.

## 1.7 Organization of the Report

A brief introduction to sensitivity analysis methods is given in Chapter 2. The report is divided into two main parts corresponding to the *E. coli* and *Listeria monocytogenes* models. Part A includes Chapters 3 to 11 and Part B includes Chapters 12 to 20. Chapter 3 explains the *E. coli* model and presents case scenarios and modifications performed in the original model. Chapters 4 to 9 present the results of nominal range sensitivity analysis, analysis of variance, regression analysis, classification and regression tree, scatter plots and conditional sensitivity analysis, respectively in the *E. coli* model. Chapter 10 presents the results of exposure assessment in ground beef servings. Chapter 11 summarizes the conclusions and recommendations based on the analyses of the *E. coli* model in Chapters 4 to 10.

Chapter 12 explains the *Listeria monocytogenes* model and presents the case scenarios and modifications performed in the original model. Chapters 13 to 18 present the results of nominal range sensitivity analysis, differential sensitivity analysis, regression analysis, analysis of variance, classification and regression tree, scatter plots and conditional sensitivity analysis, respectively. Chapter 19 summarizes the conclusions and recommendations based on the analyses of the *Listeria monocytogenes* model. Chapter 20 answers the questions raised in Section 1.1 based on the results of application of sensitivity analysis to both the *E. coli* and *Listeria monocytogenes* models.

# 2 SENSITIVITY ANALYSIS METHODS

The objective of this chapter is to briefly review typical sensitivity analysis methods and to recommend the selection of methods to apply to one or both of the case study food safety risk assessment models. Sensitivity analysis methods may be broadly classified as mathematical methods, statistical (or probabilistic) methods, and graphical methods. This classification helps in understanding applicability of sensitivity analysis methods for different types of models, and in selecting appropriate methods according to their usefulness to a decision-maker. Mathematical methods are useful for deterministic and probabilistic models. Statistical methods are generally used for probabilistic models. Graphical methods are usually complimentary to mathematical and statistical methods. Graphical methods can be used for any kind of model (Frey and Patil, 2002). Specific methods in each of these three categories are reviewed. Methods selected for case studies and evaluations are identified.

## 2.1 Mathematical Methods for Sensitivity Analysis

Mathematical methods assess sensitivity of a model output to the range of variation of an input. These methods typically involve calculating the output for a few values of an input within the possible range (e.g. Salehi *et al.*, 2000). For example, the output of a model can be calculated for the highest and lowest possible values of an input. Sensitivity is usually described in terms of relative change in the output. These methods do not address the variance in the output due to the variance in the inputs, but they assess the impact of range of variation in the input values on the output (Morgan and Henrion, 1990). Mathematical methods are helpful in screening the most important inputs (e.g., Brun *et al.*, 1997). Mathematical methods can be used to identify inputs that require further data identification and research in the case of deterministic models (e.g., Ariens *el al.*, 2000).

Frey and Patil (2002) discussed four methods for mathematical sensitivity analysis, including nominal range sensitivity analysis (NRSA), difference in log odd ratio (ΔLOR), break-even analysis, and differential sensitivity analysis (DSA) technique. NRSA and DSA were selected for application to the *E. coli* and *Listeria monocytogenes* models. ΔLOR and breakeven analysis were considered but not selected. ΔLOR requires the model output to be in the form of probability. Since neither the *E. coli* nor *Listeria monocytogenes* model has such an output, ΔLOR was not selected. Breakeven analysis requires the output to be characterized as acceptable

or unacceptable. Identification of acceptable or unacceptable risk is not explicitly performed in the E. coli or *Listeria monocytogenes* models and hence breakeven analysis was not selected.

The following sections explain the selected methods briefly. Section 2.1.1 describes NRSA and Section 2.1.2 describes the DSA technique. The description of methodology, advantages and disadvantages are covered in each section.

### 2.1.1 Nominal Range Sensitivity Analysis Method

NRSA is also 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 to propagate through a model in a probabilistic framework (Cullen and Frey, 1999). NRSA can be used to prioritize data collection needs as demonstrated by Salehi et al. (2000).

#### 2.1.1.1 Description

NRSA is used to evaluate the effect on model outputs 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 magnitude of the input variable is referred to as the sensitivity or swing weight of the model to that particular input variable (Morgan and Henrion, 1990). The sensitivity also can be represented as a positive or negative percentage change compared to the nominal solution. The sensitivity analysis can be repeated for any number of individual model inputs. The sensitivity index is calculated as follows:

$$\text{Sensitivity} = \frac{\text{Output}_{\text{max input}} - \text{Ouput}_{\text{min input}}}{\text{Output}_{\text{nominal input}}} \qquad (2\text{-}1)$$

The results of NRSA are most valid when applied to a linear model. In such cases, it would be possible to rank order the relative importance of each input based upon the magnitude of the calculated sensitivity measure as long as the ranges assigned to each sensitive input are accurate. However, for a non-linear model, the sensitivity of the output to a given input may depend on interactions with other inputs, which are not considered. Thus, the results of NRSA are potentially misleading for nonlinear models. In such cases, conditional NRSA can be done, in which NRSA is applied to different combinations of input values.

#### 2.1.1.2 Advantages

NRSA is a relatively simple method that is easily applied. It works well with linear models and when the analyst has a good idea of plausible ranges that can be assigned to each selected input. The results of this approach can be used to rank order key inputs only if there are no significant interactions among the inputs, and if ranges are properly specified for each input.

#### 2.1.1.3 Disadvantages

NRSA addresses only a potentially small portion of the possible space of input values, because interactions among inputs are difficult to capture (Cullen and Frey, 1999). Conditional sensitivity analysis, as described in Section 2.3.2, may be used to account for correlation between inputs or nonlinear interactions in model response, but it has limitations because of the combinational explosion of possible cases. Potentially important combined effects on the decision (or output) due to simultaneous changes in a few or all inputs together are not shown by nominal sensitivity analysis for other than linear models; thus for nonlinear models it is not clear that NRSA will provide a reliable rank ordering of key inputs.

### 2.1.2 Differential Sensitivity Analysis (DSA)

Differential Sensitivity Analysis (DSA) is a local sensitivity analysis method. It is most applicable for calculating the sensitivity of the output to small deviations in the point estimate of an input.

#### 2.1.2.1 Description

In DSA the local sensitivity is calculated at one or more points in the parameter space of an input keeping other inputs fixed. The sensitivity index is calculated based on a finite difference method. DSA is performed with respect to some point x in the domain of the model. A small perturbation $\Delta x$ with respect to the point value of a model input, such as a change of plus or minus one percent, can be used to evaluate the corresponding change in the model output. Thus, the sensitivity index may be calculated as:

$$\text{Sensitivity} = \frac{\text{Output}_{x+\Delta x} - \text{Ouput}_{x-\Delta x}}{\text{Output}_{x}} \tag{2-2}$$

A more generalized form of DSA is the Automatic Differential (AD) sensitivity analysis. AD is an automated procedure for calculating local sensitivities for large models (Grievank, 2000). In AD the local sensitivity is calculated at one or more points in the parameter space of

the model. At each point, the partial derivative of the model output with respect to a selected number of inputs is evaluated. The values of partial derivatives are a measure of local sensitivity. Automatic differentiation has been applied to models that involve complex numerical differentiation calculations such as partial derivatives, integral equations, and mathematical series (Hwang et al., 1997).

#### 2.1.2.2 Advantages

DSA is conceptually easy to apply and needs only a small amount of computational time compared to statistical methods if sensitivity at only few points is calculated. It 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. The sensitivity thus obtained can aid in identifying the significant figures needed for the point estimates of an input. DSA provides insight into the comparative change in the output associated with an equivalent perturbation of each input.

#### 2.1.2.3 Disadvantages

DSA does not consider the possible range of values that inputs can take in calculation of sensitivity indices. Thus, no inference can be made regarding global sensitivity. DSA is based on finite difference method. AD is superior to finite difference approximations of the derivatives because numerical values of the computed derivatives are more accurate and computational effort is significantly lower (Bischof et al., 1992).

For nonlinear models, DSA does not account for interaction among inputs. Therefore, the significance of differences in sensitivity between inputs is difficult to determine making the rank ordering of key inputs potentially difficult.

## 2.2 Statistical Methods for Sensitivity Analysis

Statistical methods involve running simulations in which inputs are assigned probability distributions and assessment of the effect of variance in inputs on the output distribution (e.g. Andersson *et al.*, 2000). Depending upon the method, one or more inputs are varied at a time. Statistical methods allow one to identify the effect of simultaneous interactions among multiple inputs.

Distributions for model inputs can be propagated through the model using a variety of techniques, such as Monte Carlo simulation, Latin Hypercube sampling, and other methods

(Cullen and Frey, 1999). Sensitivity of the model to individual inputs or groups of inputs can be evaluated by variety of techniques. Statistical methods are widely used for probabilistic models as these methods can evaluate the effect of variance in the inputs on the output. A probabilistic model is itself deterministic in nature, but inputs are assigned distributions (Cullen and Frey, 1999).

Frey and Patil (2002) discuss five statistical methods for sensitivity analysis, including linear regression analysis (RA), analysis of variance (ANOVA), response surface method (RSM), Fourier Amplitude Sensitivity Test (FAST), and Mutual Information Index (MII). Other statistical-based methods for sensitivity analysis were identified during the course of this work, including sample and rank regression coefficients, rank regression, Categorical and Regression Trees (CART), and Sobol's method. Of these various methods, the following were selected as the basis for one or more case studies: RA; ANOVA; sample (Pearson) correlation coefficients; rank (Spearman) correlation coefficients; and CART. More information regarding each of these methods is given in the following subsections.

Methods not selected for case studies include RSM, FAST, Sobol's method, and MII. The rationale for not including these methods in case studies is briefly summarized here. RSM was not selected because it is not typically a sensitivity analysis method in itself; rather, it is used to simplify the original model for the purpose of facilitating application of iterative sensitivity analysis methods. RSM is similar in many respects to the regression methods, although the functional form of a typical RSM is nonlinear with interaction terms.

FAST and Sobol's methods are both variance-based methods that enable apportionment of the variance in a model output to the variance in model inputs. These two methods are potentially useful and powerful methods; however, at this time software for application of these methods was not readily available that could be appropriately interfaced with the two case study models. FAST is included as a capability of the SIMLAB software that is currently undergoing commercialization. The most readily available implementation of FAST is in a C++ based software environment (SIMLAB, 2000; Giglioni, 2001) that is not easily interfaced with the Excel-based models that are the focus of the case studies here. Software for calculation of Sobol's indices was not available. Although neither FAST nor Sobol's method is applied to case studies here, a brief description of each of these two methods is included in the following

subsections. Either or both of these methods may be useful to evaluate in the future when appropriate software for their implementation becomes available.

MII is one of the few methods that can appropriately deal with complex interactions between model inputs. However, MII requires multiple evaluations of conditional probability distributions, which requires repetitive Monte Carlo simulations. For example, Patil and Frey (2003) evaluated MII applied to a draft Vibrio parahaemolyticus food safety risk model for shellfish. MII is not available in an existing software package. Even if automated, MII would require dozens or more Monte Carlo simulations per analysis and therefore was deemed to be impractical to apply to the larger food safety risk models that are the subject of case studies here.

The methods that were applied in one or more case studies to either the *E. coli* or *Listeria monocytogenes* food safety risk assessment models are summarized in the following sections, including sample and rank correlations, linear regression, ANOVA, and CART. In addition, FAST and Sobol's method are briefly discussed to facilitate future consideration of these methods even though they were not applied here.

### 2.2.1 Sample and Rank Correlation Coefficients

The correlation coefficient is a statistic that is calculated from sample data, and it is used to estimate the corresponding population parameter *r*. Correlation coefficients measure the strength of a linear the relationship between an input and the output. A correlation exists between two variables when one of them is related to the other in some way. There are two types of correlation coefficients: parametric or Pearson and, non-parametric or Spearman.

#### 2.2.1.1 Description

Correlation coefficients can range from -1 to +1. The value of -1 represents a perfect negative correlation while a value of +1 represents a perfect positive correlation. A value of zero represents a lack of correlation (Edwards, 1976). The strength of the relationship between *x* and *y* is sometimes expressed by squaring the correlation coefficient and multiplying by 100. The resulting statistic is known as variance explained (or $R^2$). For example, a correlation of 0.5 means 25% of the variance in *y* is "explained" or predicted by the *x* variable. The correlation between two variables *x* and *y* is defined as (Steel et. al., 1997):

$$r = \frac{\sigma_{XY}}{\sigma_X \times \sigma_Y} = \frac{\sum_{i=1}^{n}(x_i - \bar{x}) \times (y_i - y)}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \times \sum_{i=1}^{n}(y_i - \bar{y})^2}} \quad (2\text{-}3)$$

If *x* and *y* are not closely related to each other, then their covariance is small and therefore their correlation is also small.

The inverse Fisher transformation is used to test the statistical significance of the correlation coefficients. The test is based on the assumption that the distribution of the residual values (i.e., the deviations from the regression line) for the dependent variable *y* follows the normal distribution, and that the variability of the residual values is the same for all values of the independent variable *x*. However, Monte Carlo studies suggest that meeting those assumptions closely is not absolutely crucial if the sample size is not very small and when the departure from normality is not very large (Steel et. al., 1997).

There are several important kinds of correlation, differing in the details of calculation. The most widely used type of correlation coefficient is Pearson r, also called the sample, linear or product moment correlation.

The Spearman correlation coefficient is non-parametric and is also referred to as a rank correlation. The spearman correlation is similar to the Pearson correlation except that it is computed from ranks. Therefore, the Spearman correlation is a measure of the stength of the monotonic relationship between two random variables, and it can account for monotonic nonlinear relationships (Kendall, 1990). Detailed discussions of the Spearman *r* statistic, its power and efficiency can be found in Gibbons (1985), Hays (1981), McNemar (1969), Siegel (1956), Siegel and Castellan (1988), Kendall (1948), Olds (1949), or Hotelling and Pabst (1936).

#### 2.2.1.2 Application

Correlation coefficients are widely used to assess sensitivity (Cullen and Frey, 1999,Borkman et. al., 1993). Commercial software packages are available to calculate correlation coefficients using simple menu driven approach. Examples of such software are @RISK© and Crystal Ball© among many others. However, most menu driven software does not allow automation when correlation coefficients have to be calculated for a large number of datasets. In such cases, the macro features of statistical software packages such as SAS© and S-PLUS™ may be used.

#### 2.2.1.3 Advantages

The Pearson correlation coefficients capture linear relationships in the model. Spearman correlation coefficients can respond to nonlinear monotonic relationships. Both correlation coefficients are relatively easy to compute, as they are readily available in many commercial software packages.

#### 2.2.1.4 Disadvantages

Correlation does not imply causation. There can be a case where a third variable is influencing the two variables with high correlation. Pearson coefficients are inaccurate for nonlinear models and Spearman coefficients are inaccurate for non-monotonic models. Neither Pearson nor Spearman coefficients capture complex dependencies nor directly deal with interactions.

### 2.2.2 Regression Analysis

Regression analysis can be employed as a probabilistic sensitivity analysis technique as demonstrated by Iman et al. (1985). Regression analysis serves three major purposes (Neter et al., 1996; Sen and Srivastava, 1990):

- Description of the relationship between input and output variables
- Control of input variables for a given value of the output variable
- Prediction of a output based on input variables

#### 2.2.2.1 Description

A mathematical relation between inputs and the output must be identified prior to regression analysis. Such relationship could be identified or hypothesized based upon inspection of scatter plots or upon understanding of the functional form of the model. Regression analysis is most properly performed when the output is a random sample. 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). Regression analysis typically involves fitting a relationship between inputs and an output such as this linear one:

$$Y_i = \beta o + \beta_1 X_{1,i} + \beta_2 X_{2,i} + \ldots + \beta_m X_{m,i} + \varepsilon_i \qquad (2\text{-}4)$$

where,

Yi = $i^{th}$ output data point for $i^{th}$ input data points

| | | |
|---|---|---|
| Xj,i | = | $i^{th}$ input data point for the $j^{th}$ input |
| $\beta$j | = | regression coefficient for the $j^{th}$ input |
| $\varepsilon_t$ | = | error for the $i^{th}$ data point |

Each term in the regression model can have a different basis function, which can be linear or nonlinear. Most typically, each basis function is a linear function of only one input. If the analyst has a priori knowledge of more appropriate functional forms, they can be used instead. For a linear model, the regression coefficient $\beta_j$, can be interpreted as the change in output $Y_i$ when the input $X_{j,i}$ for a given value of $j$ increases by one unit and the values of all other inputs remain fixed (Devore and Peck, 1996). Therefore, regression coefficients can be used as a form of nominal range sensitivity. The goodness of fit of the regression model to the data can be measured using the coefficient of multiple determinations, $R^2$. $R^2$ is a measure of the amount of variance in the dependent variable explained by the model (Draper and Smith, 1981). A key assumption of least squares regression analysis is that the residuals are normally distributed.

Because the regression coefficients are estimated from a random sample of data, the estimated regression coefficients themselves are random variables. If the coefficient is not significantly different than zero, then there is not a statistically significant linear relationship between the input and the output (Draper and Smith, 1981). Conversely, if the coefficient is statistically significant, then there is stronger evidence of sensitivity. To determine statistical significance, the standard error of the regression coefficient is estimated. If the ratio of the value of the regression coefficient divided by its standard error is greater than a critical value, then the coefficient is deemed to be statistically significant. The critical value is determined based upon the desired significance level (usually 0.05) and the degrees of freedom of the regression model (Devore and Peck, 1996). The magnitude of statistically significant regression coefficients can be used to help determine the ranking of the inputs according to their sensitivity if the inputs or the coefficients are normalized (or standardized) to remove dimensional effects (Neter et al., 1996; Iman et al., 1985).

In regression analysis, in order to evaluate the possibility of discarding insignificant explanatory variables from the model, a specific statistical test, called "Full versus Reduced F-test" should be performed (Steel and Dickey 1996). The null hypothesis in this test is presented in Equation 2-5:

$$H_0 = \beta_m = \beta_n = \Lambda = \beta_t = 0 \quad \text{versus} \quad \text{Not So} \tag{2-5}$$

The null hypothesis tests if all the coefficients are simultaneously zero. If this hypothesis is satisfied then none of the inputs corresponding to these coefficients have a significant effect on the output of the model. A full model considers all inputs, whereas a reduced model considers only significant inputs. Referring to full and reduced models, the following ratio should be calculated. $H_0$ should be rejected if $F_{Calc} > F_{Critical}$:

$$F_{Calc} = \frac{\frac{SSM_{Full} - SSM_{Reduced}}{dfM_{Full} - dfM_{Reduced}}}{MSE_{Full}} \tag{2-6}$$

where,

SSM = Model sum of squares

dfM = Model degrees of freedom

MSE = Error mean square

In order to compare input variables on the basis of regression coefficients, the data must be normalized. However, instead of the typical standardization approach, in which subtracting the mean and dividing by the standard deviation normalize the data points, the normalized value is also divided by a factor accounting for the sample size. The following equation is used for normalization (Neter et al., 1996).

$$X' = \frac{1}{\sqrt{n-1}}\left(\frac{X - \overline{X}}{\sigma}\right) \tag{2-7}$$

where,

X' = Normalized data point;

n = Number of data points

$\overline{X}$ = Mean of the data set

σ = Standard deviation of the data set

Incorporating the detail formulation for standard deviation, the transformed variable, X' can be represented as:

$$X' = \frac{1}{\sqrt{n-1}} \times \frac{X - \overline{X}}{\frac{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}}{\sqrt{n-1}}} = \frac{X - \overline{X}}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2}} \leq \pm 1 \tag{2-8}$$

Therefore, the transformed variable will always lie between +1 and -1. Thus, regression coefficients can be compared and expressed in terms of percentages.

Regression analysis can handle both qualitative and quantitative inputs. Quantitative inputs take on values on a well-defined scale: examples are storage time, storage temperature, and decontamination efficiency. In contrast, many inputs of interest are not quantitative but are qualitative. Examples of qualitative inputs are gender (male, female), season (summer, winter), and ground beef consumption type (hamburger, raw, meatballs). The method used for addressing qualitative inputs in regression analysis is explained here. This method is used in case studies for the *E. coli* model in which there are qualitative inputs.

There are many ways of identifying the levels of a qualitative input. One of these methods is application of indicator variables that can take on values of 0 and 1 (Neter *et al.,* 1989). The use of indicator variables for each level of a qualitative input leads to computational difficulties such as singularity in the matrix of coefficients. An approach that can be used to avoid this difficulty is to add an extra condition to the model by dropping one of the indicator variables and assuming that the summation of level effects is zero. Hence, a qualitative input with $c$ levels is represented by $c - 1$ indicator variables, each taking on the values 0 and 1. Thus, if all of the $c - 1$ indicators have a value of 0, it is implied that the $c^{th}$ level is chosen. Any solution to the regression model with a qualitative input will provide estimated coefficients for each of the indicator variables and not for the qualitative input directly. Hence, it is important to understand the meaning of the regression coefficients when there is a qualitative input in the model. If there is no interaction between the qualitative and quantitative inputs to the model, estimated coefficients for the indicator variables have the effect of adjusting the intercept of the regression model conditional on the selected level of the qualitative input. This causes the fitted regression model to shift up or down based on the magnitude of the coefficients.

Regression analysis is used as a method for evaluating the sensitivity of the output to the inputs of the model using the estimated regression coefficients of quantitative inputs on a comparative basis. For a qualitative input, there are typically multiple coefficients estimated for the indicator variables and not one coefficient for the original qualitative input. Furthermore, because the indicator variables can have values of only 0 and 1, the magnitude of their coefficients does not have the same interpretation as that for a quantitative input. Moreover, estimated coefficients for indicator variables are not unique. These coefficients depend on the assumption made for solving the singularity problem (e.g., sum of level effects is zero).

Therefore, the estimated coefficients for the indicator variables cannot be used in a comparative basis with those of the quantitative inputs.

As an alternative approach for sensitivity analysis when there is a qualitative input to the model, F values estimated for both qualitative and quantitative inputs are used as indices of sensitivity. An F value corresponding to each input represents the ratio of the mean input sum of squares to the mean error sum of squares (Neter *et al.,* 1996). A statistically significant F value indicates that there is a statistically significant effect corresponding to the input and the input cannot be discarded from the regression model. Moreover, for quantitative inputs it can be proved that the F value is equivalent to the square of the t value. The t value is the ratio of the estimated coefficient for the quantitative input to the standard error corresponding to the estimated coefficient. Large F values correspond with cases where the standard error for the input is small indicating that there is a small amount of uncertainty regarding estimated coefficient. In those cases, the range of the estimated confidence interval for the coefficient is narrow. In contrast, small F values correspond with cases where the standard error for the estimated coefficient is large indicating that there is large amount of uncertainty regarding estimated coefficient. In these latter cases, the range of the estimated confidence interval for the coefficient is wide. Hence, F values not only take into account the magnitude of the coefficient, but also consider the amount of error corresponding to each coefficient.

Regression analysis can be applied using commercial software software's like SAS© and S-PLUS™. Macro feature in these software packages is used to automate the analysis when applied to large number of datasets.

#### 2.2.2.2 Advantages

Regression techniques such as the ones discussed here allow evaluation of sensitivity of individual model inputs, taking into account the simultaneous impact of other model inputs on the result (Cullen and Frey, 1999). Other regression techniques, such as those based upon the use of partial correlation coefficients, can evaluate the unique contribution of a model input with respect to variation in a selected model output (Brikes and Dodge, 1993). Moreover, a rank regression approach may also be used. In rank regression, the ranks of each input and the output are used instead of the sample values. Rank regression can capture any monotonic relationship between an input and the output, even if the relationship is nonlinear. Sample and rank regression methods are discussed elsewhere, such as by Neter *et al.* (1996), Iman *et al.* (1985),

Brikes and Dodge (1993), and Kendall and Gibbons (1990). Also, methods such as the change-point regression that estimate the parameters corresponding to the points at which the slopes change may be able to identify thresholds in the model (Ogden and Parzen, 1996).

#### 2.2.2.3 Disadvantages

The key potential drawbacks of regression analysis include: possible lack of robustness if key assumptions of regression are not met; the need to assume a functional form for the relationship between an output and selected inputs; and potential ambiguities in interpretation. Regression analysis works best if each input is statistically independent of every other input (Devore and Peck, 1996). Furthermore, the residuals of a least squares regression analysis must be normally distributed and independent. If these conditions are violated, the results of the analysis may not have a strict quantitative interpretation, but instead should be treated as providing conceptual or qualitative insights regarding possible relationships. The results of regression analysis can be critically dependent upon the selection of a functional form for the regression model. Thus, any results obtained are conditioned upon the actual model used. Regression analysis can yield results that may be statistically insignificant or counter-intuitive (Neter et al., 1996). The lack of a clear finding may be because the range of variation of that input was not wide enough to generate a significant response in the output. Thus, regressions results can be sensitive to the range of variation in the data used to fit the model and may not always clearly reveal a relationship that actually exists. The regression model may not be useful when extrapolating beyond the range of values used for each input when fitting the model (Devore and Peck, 1996).

### 2.2.3 Rank Regression

Rank regression is a regression method where input and output values are rank ordered and the linear association between the ranks of an output and corresponding inputs is estimated in terms of rank regression coefficients (Neter et. al., 1996).

#### 2.2.3.1 Description

The procedure for rank order regression is similar to that of stepwise linear regression except that ranks are used instead of sample values. The input and output values are rank ordered. A regression model, minimizing sum of squares for the output, is fit to the ranked data.

A high $R^2$ value indicates a monotonic relationship. The rank regression coefficient can be used to rank the inputs.

#### 2.2.3.2 Application

Sample and rank regression methods are discussed elsewhere, such as by Neter *et al.* (1996), Iman *et al.* (1985), Brikes and Dodge (1993), and Kendall and Gibbons (1990). Rank regression can be applied using commercial software such as SAS© and S-PLUS™. The macro features of the software can be used for automation when applied to a large number of datasets.

#### 2.2.3.3 Advantages

Rank regression can capture any monotonic relationship between an input and the output, even if the relationship is nonlinear. Rank regression is especially useful when there is high amount of variance or noise in the data (Steel, et. al., 1997). Rank regression can be computationally more efficient as it does not have to deal with large numbers, especially outliers and decimal digits in the original data. Instead, all inputs have the same uniformly distributed range of ranks from 1 to n, where n is the sample size.

#### 2.2.3.4 Disadvantages

Rank regression assumes a monotonic model and thus is not applicable for models explained by non-monotonic functions. Rank regression coefficients, unlike standard regression coefficients, cannot be transformed to obtain sensitivities in terms of the original ranges of each input. Non-linearity in the response cannot be directly inferred from rank regression results.

### 2.2.4 Analysis of Variance

ANOVA is a probabilistic sensitivity analysis method used for determining whether there is a statistical association between an output and one or more inputs (Krishnaiah, 1981). ANOVA differs from regression analysis in that regression analysis is used to form a predictive model whereas ANOVA is a general technique that can be used to test the hypothesis that the means among two or more groups are equal, under the assumption that the mean of the outputs for each of the groups is normally distributed with same variance (Neter et. al., 1996). Also, ANOVA addresses both categorical inputs and groups of inputs (Steel et. al., 1997).

#### 2.2.4.1 Description

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" and a "contrast" is a linear combination of two or more factor level means. For example, a contrast can be built to evaluate the mean growth of pathogens when the storage temperature varies between high and low levels for a specific storage time. Discrete variables are easily treated as levels. The continuous variables can be partitioned to create levels. For example, storage temperature in the *Listeria monocytogenes* model is a continuous factor that can take any value between 0 $^0$C to 10 $^0$C. In order to define levels for this factor, this range can be divided into two discrete levels representing low and high temperatures. Temperature values between 0 $^0$C and 5 $^0$C are considered for the low level and temperatures between 5 $^0$C to 10 $^0$C are considered for the high level. Single-factor ANOVA is used to study the effect of one factor on the response variable. Multifactor ANOVA deals with two or more factors and it is used to determine the effect of interactions between factors. A qualitative factor is one where the levels differ by some qualitative attribute, such as a type of pathogen or geographic regions (Neter et al., 1996).

ANOVA is used to determine if the mean values of the output vary in a statistically significant manner associated with variation in values for one or more inputs. If the mean response of the output does not have a significant association with variation in the inputs, then the variation in the output is random. ANOVA uses the F test to determine whether there exists a significant difference among treatment means or interactions. If the null hypothesis (no difference among treatments or interactions) is accepted, there is an implication that no relation exists between the factor levels and the response. When the F test rejects the null hypothesis, thorough analysis of the nature of the factor-level effects should be undertaken (Neter et al., 1996).

In ANOVA, it is assumed that the replications for a treatment are done by sampling from a population normal distribution. These population normal distributions corresponding to each treatment are assumed to have the same variance but different population means. As a result of these assumptions, the mean of responses for a treatment is also a random sample from a normal distribution whose variance is population variance for the treatment divided by number of replications. Also, the mean of this normal distribution is the same as the population mean for the normal distribution corresponding to the treatment (Neter et. al., 1996). Diagnostic checks are important to determine whether the assumptions of ANOVA are violated. If any key assumptions are violated then there can be corrective measures to address the problem. The F

test is generally robust to deviations from these assumptions but substantial departures from normality or large differences in the variances of the output can influence statistical test results (Lindman, 1974). In the case of correlated inputs, the results of the F test may not be robust. However, approaches such as principal component analysis to group correlated factors can be used to address this problem (Kim and Mueller, 1978).

In ANOVA, the statistical significance of factors is tested based on F values. The F values can be used to rank the factors based on their relative magnitude. The higher the F value for a factor, the more sensitive is the output to the factor. Therefore, factors with higher F values are given higher rankings. The sum of squares for each factor may be considered as an alternative measure of sensitivity. However, F value is preferred as it accounts for not only the sum of squares but also the degree of freedom associated with the factor (Carlucci, 1999).

The F values calculated for each effect indicate the statistical significance of the respective effect. The $R^2$ value for ANOVA indicates the fraction of output variance captured by the main and interaction effects considered in the model. Moreover, a high $R^2$ implies that results are not compromised by inappropriate definition of the levels for each factor. Thus, the $R^2$ can be used as diagnostic for ANOVA.

Commercial software's such as SAS© and S-PLUS™ allow application of ANOVA. When ANOVA is applied to a large number of datasets macro feature of these software's is used to automate the application and summarization of results.

#### 2.2.4.2 Advantages

ANOVA can be used to analyze both continuous and discrete factors (Montgomery, 1997). The results of ANOVA can be robust to departures from key assumptions, and additional techniques can be employed to deal with issues such as multi co-linearity. ANOVA allows evaluation of the "main effect" between factors. The main effect is the effect of the factor alone, averaged across the levels of other factors. ANOVA can also be used to evaluate the "interaction effect" between factors. An interaction effect is the variation among the differences between means for different levels of one factor over different levels of another factor. For example, when the difference between the mean responses for high and low levels of the storage temperature at a low level of the storage time is not equal to difference between the mean responses for high and low levels of the storage temperature at a high level of the storage time, there is an interaction between the storage time and the storage temperature. If there is a

significant interaction, detailed contrasts can be evaluated. By comparing results for different levels of each factor, it might be possible to identify thresholds in the model response.

#### 2.2.4.3 Disadvantages

ANOVA can become computationally intensive if there are a large number of inputs. If this becomes a problem, a suggestion by Winter et al. (1991) is to try to reduce the number of inputs analyzed by using some less computationally intensive method, such as NRSA, to screen out insensitive inputs. If there is a significant departure of the response variable from the assumption of normality, then the results may not be robust (Lindman, 1974). Errors in the response variables due to measurement errors in the inputs can result in biased estimates of the effects of factors. If the inputs are correlated, then the effect of each individual input on the response variable can be difficult to assess (Neter et al., 1996), unless methods such as principal component analysis are used.

In unbalanced experiments with unequal numbers of observations in different treatments not all contrasts may be estimable. A contrast is not estimable if the variables involved are not independent but depend upon a combination of other variables **(**Giesbrecht, and Gumpertz**,** 1996).

### 2.2.5 Classification and Regression Tree

CART or hierarchical tree-based regression (HBTR) can be thought of as a forward stepwise variable selection method, analogous to forward stepwise regression analysis. The method used to estimate regression trees has been around since the early 1960's. The method proceeds by iteratively asking and answering following questions (Breiman et al., 1984):

- Which variable of all independent variables 'offered' in the model should be selected to produce the maximum reduction in variability of the dependent variable (response)?
- Which value of the selected variable (discrete or interval) results in the maximum reduction in variability of the response?

Numerical search procedures are applied iteratively until a desirable end-condition is met, at which time the final tree structure is formed. The CART terminology is similar to that of a tree; there are branches, branch splits or internal nodes, and leaf or terminal nodes (Washington et al., 1997). The components of the classification and regression tree are shown in Figure 2-1.

A node is an input variable based on which data is split. Nodes can be a root node, intermediate nodes or leaf nodes. A root node is the node based on which the data is first split.

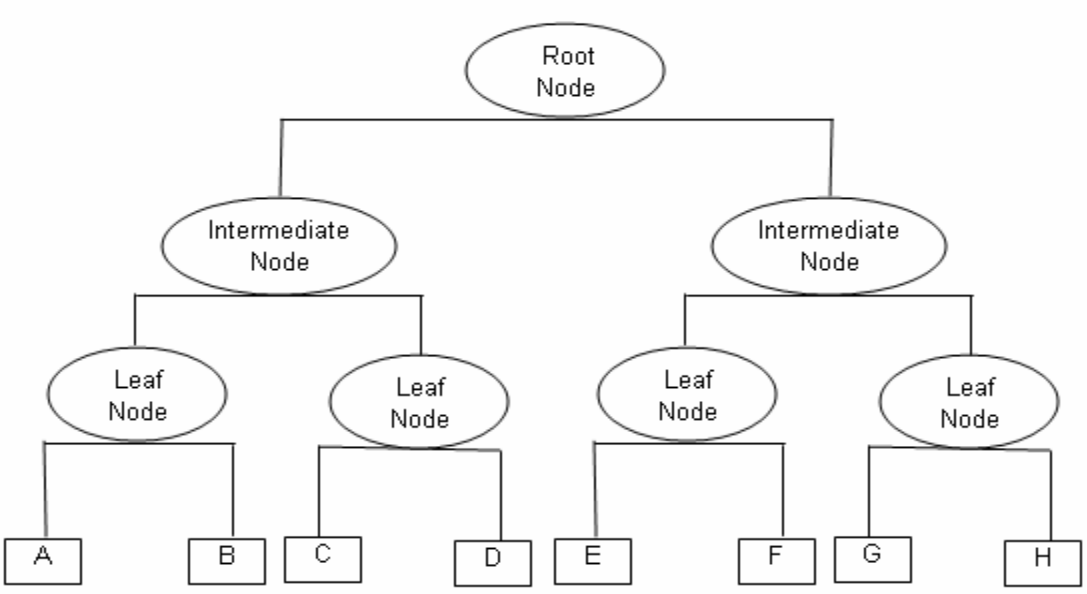



Figure 2-1. Schematic Diagram of a Classification and Regression Tree Illustrating Rout Node, Intermediate Nodes, and Terminal Leaves.

Intermediate nodes are the nodes on the basis of which the data is successively split. Leaf nodes are the nodes on which the penultimate data was split. Branches are the conditions on the input variables that determine which input set goes to which new dataset. A set of conditions on the input variable from the root node leading to a root node is called a path or classification rule.

2.2.5.1 Description

CART conceptually seeks to divide a data set into subsets, each of which is more homogeneous compared to the total dataset. At a given level of division, each of the subsets is intended to be different in terms of the mean value. Thus, CART is a statistical approach for binning data.

In order to explain the method in mathematical terms, the definitions presented by Washington et al. (1997) are used. The first step is to define the deviance at a node. A node represents a data set containing L observations. The deviance, $D_a$, can be estimated as follows:

$$D_a = \sum_{l=1}^{L} \left( y_{l,a} - \bar{y}_a \right)^2 \qquad (2\text{-}9)$$

where,

$D_a$ = total deviance at node a, or the sum of squared error at the node

$y_{l,a}$ = $l^{th}$ observation of dependent variable y at node a

$\bar{y}_a$ = estimated mean of L observations in node a

For each of k variables, the algorithm seeks to split the domain of a variable, $X_i$, (where i has a value from 1 to k) into two half-ranges at node a, resulting in two branches and corresponding nodes b and c, each containing M and N of the original L observations (M + N = L) of the variable $X_i$. The reduction in deviance function is then defined as follows:

$$\Delta_{(allX)} = D_a - D_b - D_c \tag{2-10}$$

where:

$\Delta_{(allX)}$ = the total deviance reduction function evaluated over the domain of all $X_i$'s (i.e. for k number of X variables)

$D_b$ = $\sum_{m=1}^{M} \left(y_{m,b} - \bar{y}_b\right)^2$

$D_c$ = $\sum_{n=1}^{N} \left(y_{n,c} - \bar{y}_c\right)^2$

$D_b$ = total deviance at node b

$D_c$ = total deviance at node c

$y_{m,b}$ = $m^{th}$ observation of dependent variable y in node b

$y_{n,c}$ = $n^{th}$ observation of dependent variable y in node c

$\bar{y}_b$ = estimated mean of M observations in node b

$\bar{y}_c$ = estimated mean of N observations in node c.

The method seeks to find $X_k$ and its optimal split at a specific value of $X_k$, $X_k(i)$, so that the reduction in deviance is maximized. The maximum reduction occurs at a specific value $X_k(i)$, of the independent variable $X_k$. When the data are split at $X_k(i)$, the remaining samples have a smaller variance than the original data set. Numerical methods are used to maximize (Equation 2-10) by varying the selection of which variable to use as a basis for a split and what value to use at the split point. The iterative partitioning process is continued at each node until one of the following conditions is met: (1) the node of a tree has met minimum population criteria which is the minimum sample size at which the last split is performed; or (2) minimum deviance criteria at a node have been met. Some software, such as S-PLUS™, allows the user to select either criterion.

Although it might be possible that several inputs affect the response of the model, CART considers only those inputs having a significant effect on the variability of the response. The

reduction in deviance associated with the inputs present in the tree can be used as sensitivity index to rank the inputs. CART also provides an indication of priority among different inputs based upon their precedence in the tree. Typically those inputs in the top nodes have more importance and influence on the response variable in comparison with inputs in the lower nodes. Furthermore, it is possible that an input will be selected repeatedly for multiple levels within the tree, which is also an indication of the importance of that input. If there is a threshold, it is picked up at the splitting point. However, there is no guarantee that a particular splitting point is also a threshold.

#### 2.2.5.2 Advantages

One of the advantages of CART over traditional regression analysis is that it is a nonparametric method and does not require assumptions of a particular distribution for the error term or of a functional form for the relationship between the input and the output. CART is more resistant to the effects of outliers since splits usually occur at non-outlier values (Roberts et al., 1999). A regression tree selects only the most important independent variables and values of these variables that result in the maximum reduction in deviance. Results are invariant with respect to monotonic transformations of the independent variables. As a result the researcher does not have to test a number of transformations to find the "best "fit (Hallmark et al., 2002). Moreover, application to discrete and continuous explanatory variables and also qualitative variables is possible in the CART method.

#### 2.2.5.3 Disadvantages

At times, difficulty in prioritizing the explanatory variables based on the results of the CART method can be considered as a disadvantage of the method. The input variable at the first splitting point is often the most important variable among others, but in lower branches it is not always possible to easily compare variables with regard to their importance. These points are illustrated in the case studies of Chapters 7 and 17 for the *E. coli* and *Listeria monocytogenes* models, respectively. Moreover, a method for evaluating sensitivity based upon the contribution of each variable to the total reduction in deviance is illustrated in Section 7.3.1.1 and Section 17.1.1.

### 2.2.6 Sobol's Indices

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 A is given by S(A) + S(AB) + S(ABC), where S(x) is the sensitivity index of x. Unlike Sobol's methods, methods that involve only correlation ratios (Kendall and Stuart 1979; Krzykacz, 1990) or importance measures (Hora and Iman, 1990) consider just the main effect of an input, and do not account for the effect of interactions among two or more inputs.

#### 2.2.6.1 Description

The underlying principle upon which Sobol's approach calculates the sensitivity indices is the decomposition of function f(x) into summands of increasing dimensionality (Chan *et al*., 2000):

$$f(x_1,.....x_n) = f_0 + \sum_{i=1}^{n} f_i(x_i) + \sum_{i=1}^{n}\sum_{j=i+1}^{n} f_{ij}(x_i, x_j) + ..... + f_{1,2,.....,n}(x_1,....,x_n) \quad (2\text{-}11)$$

The form presented in Equation 2-11 can only be arrived at when $f_0$ is a constant, and the integral of every summand over any of its own variables is always zero, i.e.

$$\int_0^1 f_{i_1,.....,i_s}(x_{i_1},.....,x_{i_s})dx_{i_k} = 0, if 1 \le k \le s \quad (2\text{-}12)$$

Where,

s = summand index

k = input variable index

A consequence of Equation 2-11 and 2-12 is that all the summands in Equation 2-11 are orthogonal, i.e. if $(i_1,....., i_s) \neq (j_1,....., j_l)$, then

$$\int_{K^n} f_{i_1,\ldots,i_s} f_{j_1,\ldots,j_l}\, dx = 0 \tag{2-13}$$

Where, $K^n$ is the n-dimensional space of input parameters. The total variance D of f(x) is defined to be

$$D = \int f^2(x)dx - f_0{}^2 \tag{2-14}$$

and the partial variances are computed from each of the terms in Equation 2-11.

$$D_{i_1,\ldots,i_s} = \int_0^1 \ldots \int_0^1 f_{i_1,\ldots,i_s}(x_1,\ldots,x_s)dx_{i_1}\ldots dx_{i_s} \tag{2-15}$$

Where $1 \le i_1 < \ldots < i_s \le n$ and s=1,.., n. By squaring and integrating Equation 2-11 over $K^n$, and by Equation 2-13 we have

$$D = \sum_{i=1}^{n} D_i + \sum_{i=1}^{n}\sum_{j=i+1}^{n} D_{ij} + \ldots + D_{1,2,\ldots n} \tag{2-16}$$

Thus, a sensitivity measure $S(i_1, \ldots, i_s)$ is defined as

$$S(i_1,\ldots,i_s) = \frac{D_{i_1,\ldots,i_s}}{D} \tag{2-17}$$

The sum of all the sensitivity indices is always unity. The integrals in Equation 2-14 and 2-15 can be computed by the Monte Carlo (MC) integral method.

#### 2.2.6.2 Application

The use of Sobol's indices in the field of sensitivity analysis is new and there are few publications on application of global sensitivity methods using Sobol's indices. In the field of environmental risk assessment, use of sensitivity indices such as Sobol's indices in calibration and reduction of models have been demonstrated on eutrophication models (Ratto *et al*., 2001a).

An application to a very simple chemical system consisting of the observation of the time evolution of an isothermal first order irreversible reaction in a batch system is presented in Ratto *et al*. (2001b). Sobol's method of sensitivity analysis has been used in the field of financial risk to identify the major sources of error among the several factors involved in 'the delta-hedging problem' (Compolongo, 2002).

Effort has been made to reduce the computational complexity associated with calculation of Sobol's indices. Saltelli (2002a) discusses how to make best use of model evaluation to calculate Sobol's sensitivity indices.

#### 2.2.6.3 Advantages

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 capture 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.

#### 2.2.6.4 Disadvantages

Sobol's method is a global method of sensitivity analysis. Global methods are based on the sampling of the distribution function of the input factors and on the repeated execution of the model, in order to determine the distribution of the output; therefore they are, in general, computationally intensive (Pastres et. al., 1999). Also, the ease of application depends on the complexity of the model. Hence, it is difficult to apply to models with large number of inputs and complex model structure such as modularity. There is no readily available software that facilitates application of Sobol's method.

Variance based methods provide a factor-based decomposition of the output variance, and implicitly assume that the second central moment is sufficient to describe output variability. However, when the region of interest is the tails of the output distribution, this assumption is not valid (Saltelli, 2002b).

### 2.2.7 Fourier Amplitude Sensitivity Test (FAST)

The Fourier Amplitude Sensitivity Test (FAST) method is a procedure that can be used for both uncertainty and sensitivity analysis (Cukier *et al.*, 1973, 1975, and 1978). The FAST method is used to estimate the expected value and variance of the output, and the contribution of

individual inputs to the variance of the output (Cukier *et al.*, 1973). The FAST method is independent of any assumptions about the model structure, and works for monotonic and non-monotonic models (Saltelli *et al*., 2000). The effect of only one input (local sensitivity) or the effect of all inputs varying together can be assessed by FAST.

#### 2.2.7.1 Description

The main feature of FAST is a pattern search method that selects points in the input parameter space, and which is reputed to be faster than the Monte Carlo method (McRae *et al.*, 1982). The classical FAST method is not efficient to use for high-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. Sobol's sensitivity method is similar to the FAST method and can account for interacting terms, but it is less efficient than extended FAST (Sobol, 1993).

A transformation function is used to convert values of each model input to values along a search curve. As part of the transformation, a frequency must be specified for each input. By using Fourier coefficients, the variance of the output is evaluated (Cukier *et al.*, 1973). The contribution of input $x_i$ to the total variance is calculated based on the Fourier coefficients, fundamental frequency $\omega_i$, and higher harmonics of the frequency as explained by Cukier *et al.* (1975). The ratio of the contribution of each input to the output variance and the total variance of the output is referred to as the first order sensitivity index and can be used to rank the inputs (Saltelli *et al.*, 2000). The first order indices correspond to the contribution of individual inputs and not to the contribution of interactions among inputs. To account for the residual variance in the output due to higher order or interaction terms that is not explained by first order indices, the extended FAST method is used (Saltelli *et al.*, 1999).

The model needs to be evaluated at a sufficient number of points in the input parameter space such that numerical integration can be used to determine the Fourier coefficients (Saltelli *et al.*, 2000). The minimum sample size required to implement FAST is approximately eight to ten times the maximum frequency used. In the case of discrete inputs, if a sufficiently large sample size is not available, then the output can have frequent discontinuities. In such a case, the Fourier coefficients may not be estimated properly and hence, the reliability of the results can be lower in the case of discrete inputs. The Sobol's method is capable of handling discrete inputs (Saltelli *et al.*, 2000).

McRae *et al.* (1982) describe mathematical basis and computer implementation of the FAST method. Cukier *et al.* (1978) and Saltelli *et al.*, (2000) give details of producing optimal frequency sets. Different search curves and their transformation functions used in FAST are given by McRae *et al.* (1982) and Cukier *et al.* (1975).

#### 2.2.7.2 Application

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).

As an example, Lu and Mohanty (2001) used the FAST method for sensitivity analysis of a model developed for performance assessment of a proposed nuclear waste repository. The model output is the amount of radiation for long time periods. Because the number of inputs of the model is too large to be handled by the FAST method, less important input parameters were first screened out. FAST was implemented using twenty inputs. For a 10,000 year time period of interest, the top three most important inputs identified using FAST were thermal conductivity of the rock material, the alluvium retardation coefficient for technetium, and the well pumping rate for the farming receptor group located at 20 km. Conditional complementary cumulative distribution functions of the model output (Mohanty and McCartin, 1998) were used to verify the ranking of the influential parameters produced by the FAST method. The ranking of top three parameters was found to be robust but the FAST method could not consistently rank other inputs of the set.

#### 2.2.7.3 Advantages

The FAST method is superior to local sensitivity analysis methods because it can apportion the output variance to the variance in the inputs. It also can be used for local sensitivity analysis with little modification (Fontaine *et al.*, 1992). It is model independent and works for monotonic and non-monotonic models (Saltelli *et al*., 2000). Furthermore, it can allow arbitrarily large variations in input parameters. Therefore, 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 runs (Saltelli *et al*., 2000). The FAST method can be used to determine the difference in sensitivities in terms of the

differing amount of variance in the output explained by each input and, thus, can be used to rank order key inputs.

2.2.7.4 Disadvantages

The FAST method 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). However, the extended FAST method developed by Saltelli *et al.*, (1999) can account for high-order interactions. The reliability of the FAST method can be poor for discrete inputs (Saltelli *et al*., 2000). Current software tools for FAST are not readily amenable to application to the selected food safety risk assessment models.

## 2.3 Graphical Methods for Sensitivity Analysis

Graphical methods give representation of 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). Graphical methods can be used as a screening method before further analysis of a model or to represent complex dependencies between inputs and outputs (e.g., McCamly and Rudel, 1995). Graphical methods can be used to complement the results of mathematical and statistical methods for better interpretation (e.g., Stiber et al., 1999; Critchfield and Willard, 1986).

Frey and Patil (2002) demonstrated scatter plots as an approach in graphical sensitivity analysis. This method has been selected for application in the case studies of this report. In addition a graphical method for conditional sensitivity analysis is introduced and is used in this work.

### 2.3.1 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 a method often used after a probabilistic simulation of the model. Scatter plots are also often used as a first step in other analyses such as regression analysis and response surface methods.

2.3.1.1 Description

Each realization in a probabilistic simulation (e.g. variability and uncertainty simulation), such as a Monte Carlo simulation, generates a pair of an input value and the corresponding

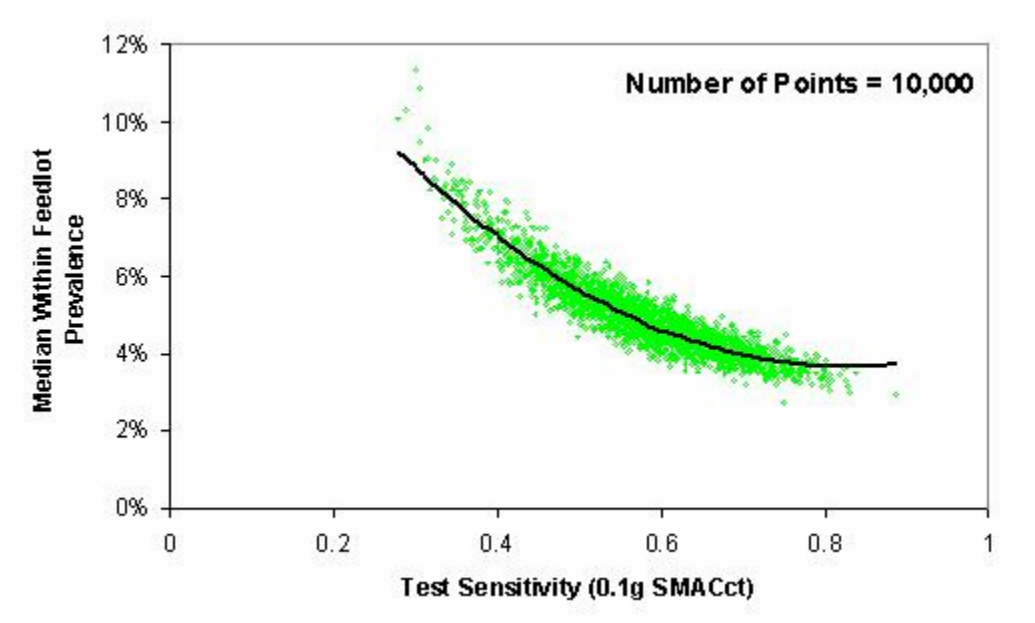



Figure 2-2. Example of a Pattern for a Scatter Plot.

output value. These simulated pairs can be plotted as points on a scatter plot. Scatter plots also can be plotted for empirical data. The scatter plot displays a range of values for both the input and output, and the general trend between them. However when data points overlap on the graph, it can be difficult or impossible to evaluate the relative frequency of occurrence of specific combinations of inputs and output values.

For example, Figure 2-1 shows simulation data for the median within feedlot prevalence of *E. coli* in summer versus the test sensitivity for the '0.1g, SMACct' testing method. This specific example is discussed in more detail in Section 8.1.2. The scatter plot was used to assess possible trends in the data and to aid in selecting a functional form for a regression model to fit to the data. In this case, there appears to be a nonlinear variation of within feedlot prevalence with respect to test sensitivity. Therefore, a nonlinear polynomial functional form was selected and fit to the data. The regression model is shown as a solid line. The comparison of the fitted model to the data is a means for verifying the adequacy of the model. It happens that in this case that the model adequately captures the key trends in the data.

#### 2.3.1.2 Advantages

Scatter plots are often recommended as a first step in sensitivity analysis of a statistical sample of data, whether it is an empirical sample or the result of a probabilistic simulation. A

key advantage of scatter plots is that they allow for the identification of potentially complex dependencies. For example, Figure 2-2 displays a nonlinear decrease in the response versus the input. An understanding of the nature of the dependencies between inputs and an output can guide the selection of other appropriate sensitivity analysis methods.

#### 2.3.1.3 Disadvantages

A potential disadvantage of scatter plots is that they can be tedious to generate if one must evaluate a large number of inputs and outputs unless commercial software is used to automatically generate multiple scatter plots (e.g., SPLUS, 2000). Although not necessarily a disadvantage, the interpretation of scatter plots can be qualitative and may rely on judgment. Whether the sensitivities of two inputs differ significantly from each other cannot always be judged from their scatter plots. When the frequency of occurrence of different combinations of inputs and outputs differ largely, overlapping of data points may affect the clarity of the graph.

### 2.3.2 Conditional Sensitivity Analysis

Conditional sensitivity analysis is considered as a graphical method since the results are often presented in form of graphs. The motivation for this technique is that the effect of variation in any one variable on the output in a non-linear model cannot be adequately captured by mathematical methods like NRSA.

#### 2.3.2.1 Description

Conditional sensitivity analysis involves repeated application of a method such as NRSA. Because a key limitation of NRSA when applied to nonlinear models is that the analysis is with respect to only one combination of input values, the objective of conditional sensitivity analysis is consider more than one such combination. For example, if a model has three inputs, in NRSA the sensitivity analysis is with respect to only one point estimate for each of the three inputs. In contrast, a conditional analysis can be with respect to multiple different combinations of point values among the three inputs. The response is calculated for point values of the selected input variable at steps or randomly generated points; the idea 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 when other variables are conditioned to fixed values. The process is repeated for other values of the other input variables.

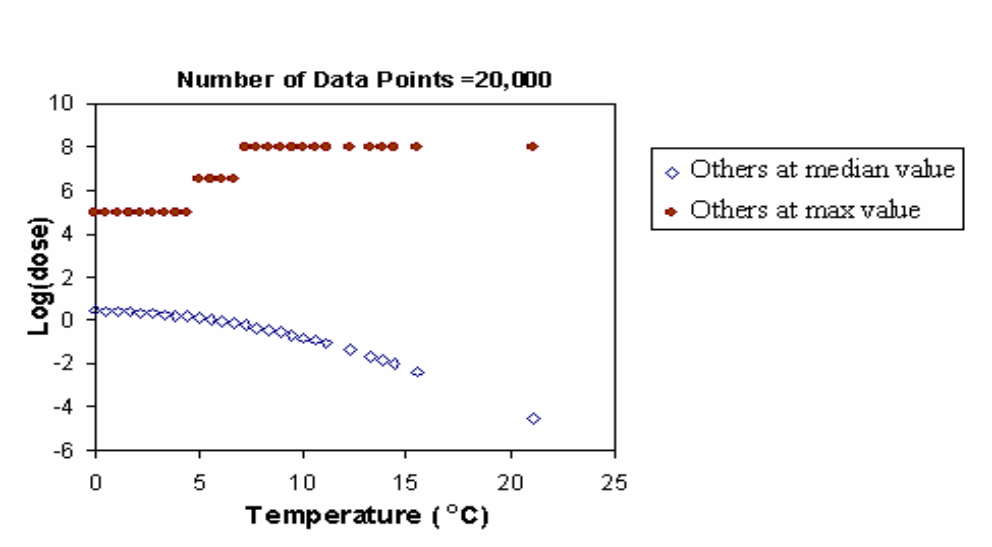



Figure 2-3. Log Dose Response for Temperature in Deli Salad.

To illustrate the methodology, exposure to *Listeria monocytogenes* is plotted versus variation in temperature in Figure 2-3. Other input variables affecting exposures were held constant at their respective median values. To generate the data points, values for temperature were randomly generated from the temperature distribution. A second set of data points were selected in which all other variables were held at their respective maximum values. The response of the exposure to temperature is shown to be different depending upon whether other inputs are held at the maximum versus median values. Thus, the response of exposure to temperature is conditional upon the values assigned to other model inputs. Similar response curves can be generated keeping other variables at minimum values. However, this response is not shown in Figure 2-3, as all dose values for the corresponding case were below the minimum detectable dose level for deli salad. Conditional sensitivity analysis enables insight into these types of interactions based upon a simple enhancement to other methods such as NRSA.

#### 2.3.2.2 Advantages

Non-linearity, saturation point and thresholds can be identified based upon conditional sensitivity analysis. These insights are under assumptions that other variables are fixed at particular values. The plotted graphs can be used to calculate NRSA and differential sensitivity indices. In many of the case studies given in Chapters 8 and 18 for the *E. coli* and *Listeria monocytogenes* models, respectively, each graph has three curves corresponding to cases where other inputs were at minimum, median and maximum values, respectively. Differential

sensitivity indices can be calculated from the curve corresponding to the case where other inputs were kept at a point value by reading the output at a point estimate, 99 percent of the point estimate and 101 percent of the point estimate. NRSA indices can be calculated by reading output values at the minimum, median and maximum values of the selected input from the curve corresponding to the case when other inputs were kept at selected point values.

#### 2.3.2.3 Disadvantages

It is not always possible to rank inputs based on the nature of the response curve alone. For example, if two inputs have response have non-linear response then it may be difficult to tell which one has higher degree of variance. This method is based on assumptions that other inputs are conditioned to specific values. Therefore a totally random outcome where multiple inputs take random values simultaneously from their distribution is not considered. Thus, this method may provide only partial insight into the nature of interactions as well as their likelihood. Hence, conditional sensitivity analysis cannot give global ranking.

## 2.4 Summary of Sensitivity Analysis Methods Selected for Case Studies

This chapter has introduced specific sensitivity analysis methods and explained which ones were selected for further evaluation. The selected methods include NRSA, AD, sample and rank correlation coefficients, RA, rank regression, ANOVA, CART, scatter plots, and conditional sensitivity analysis methods. These methods are applied to one or both of two food safety process risk assessment models as documented in later chapters.

# 3 *E. COLI O157:H7* FOOD SAFETY RISK ASSESSMENT MODEL

This chapter briefly describes the model used to estimate the occurrence of *E. coli O157:H7* in single servings of ground beef. Different modules and parts of the *E. coli* food safety risk assessment model are explained in coming sections, and specific terminologies in each module are defined in order to give better understanding of the modeling structures and the inputs and outputs in different modules. Moreover, the limitations of the original *E. coli* food safety risk assessment model with respect to the application of different sensitivity analysis methods are discussed in this chapter. Specific case scenarios for sensitivity analyses in each module are defined. Modifications performed to prepare the model for application of different sensitivity analysis methods are explained. A list of inputs and their characteristic is given for each module of the *E. coli* food safety risk assessment model.

## 3.1 Background on *E. coli O157:H7*

A German bacteriologist Dr. Theodor Escherich discovered *E. coli* bacteria in the human colon in 1885 (Riley *et. al* 1983). He showed that certain strains of the bacteria were responsible for infant diarrhea and gastroenteritis. Because scientists could grow the bacteria quickly on both simple and complex media, *E. coli* became a very popular lab organism. *E. coli* could be grown aerobically, or anaerobically. This ability classifies the *E. coli* bacteria as a facultative anaerobe. The vast majority of *E. coli* strains, including those commonly used by scientists in genetics laboratories, are harmless, however, exposure to *E. coli O157:H7* can lead to severe illness and death.

*E. coli O157:H7* infection often causes severe bloody diarrhea and abdominal cramps; which sometimes the infection causes nonbloody diarrhea or no symptoms. Usually little or no fever is present, and the illness resolves in 5 to 10 days. Certain age groups have a higher incidence of *E. coli O157:H7* infection. Surveillance from FoodNet sites in 1999 shows that 1 to 9 year-olds had the highest incidence among all age groups (CDC, 2000). In children under 5 years of age and the elderly, the infection can also cause a complication called hemolytic uremic syndrome, in which the red blood cells are destroyed and the kidneys fail. About 2%-7% of infections lead to this complication. In the United States, hemolytic uremic syndrome is the principal cause of acute kidney failure in children, and most cases of hemolytic uremic syndrome are caused by *E. coli O157:H7*.

*E. coli* is found in the family of bacteria named Enterobacteriaceae, which is informally referred to as the enteric bacteria. Other enteric bacteria are the Salmonella bacteria, Klebsiella pneumoniae, and Shigella. The latter many people consider to be a part of the *E. coli* family (Riley *et. al* 1983).

The US Department of Agriculture conducted a farm-to-table risk assessment to evaluate the public health impact from *E. coli O157:H7* in ground beef (FSIS, 2001). The risk assessment includes a comprehensive evaluation of the risk of illness from exposure to *E. coli O157:H7* in ground beef based on available data. In the risk assessment model, the likelihood of human morbidity and mortality associated with exposure to a specific number of *E. coli* pathogens consumed in ground beef is estimated. Methods to reduce the risk of illness from this pathogen in ground beef are included in the risk assessment framework (FSIS, 2001). The purpose of the risk assessment study was to:

(1) Provide a comprehensive evaluation of the risk of illness from *E. coli* in ground beef based on currently available data;
(2) Estimate the likelihood of human morbidity and mortality;
(3) Estimate the occurrence and extent of *E. coli* contamination at points along the farm-to-table continuum;
(4) Provide a tool for analyzing how to effectively mitigate the risk of illness from *E. coli* in ground beef; and
(5) Identify future food-safety research needs.

### 3.2 Overview of the Model

The *E. coli* food safety risk assessment model includes hazard identification, exposure assessment, hazard characterization, and risk characterization steps. An overview of the model is given in Figure 3-1. The hazard of *E. coli O157:H7* is identified using data from ecology, pathology, epidemiology, and microbiology. The exposure assessment consists of three major modules: (1) production; (2) slaughter; and (3) preparation. The exposure assessment is based upon a probabilistic approach for modeling the prevalence and the concentration of the *E. coli* pathogen in live cattle, carcasses, beef trim, and a single serving of cooked ground beef. In the exposure assessment several factors are taken into account, including slaughter-processing servings, consumer demographics, the consumption pattern, and seasonal differences in herd prevalence.

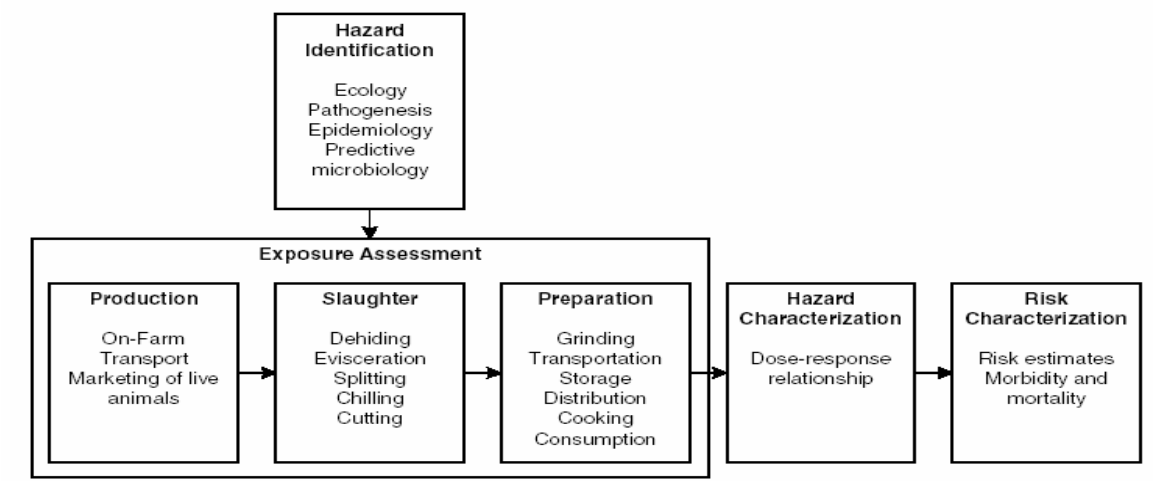



Figure 3-1. Schematic Diagram of the Farm-to-Table Risk Assessment Model for *E. coli O157:H7* in Ground Beef. Source: (FSIS, 2001).

Hazard characterization quantifies the nature and severity of the adverse effects associated with the given number of *E. coli* organisms in a ground beef serving. Risk characterization integrates the results of the exposure assessment and hazard characterization to estimate the risk of illness from *E. coli O157:H7* in ground beef. Risk estimates are provided for individuals, a community in a simulated outbreak scenario, and the U.S. population. The variability of risk among the U.S. population is considered according to differences in seasonal exposure and host susceptibility, based on the age of the consumer.

In Sections 3.2.1 to 3.2.3 different modules inside the exposure assessment part are explained. These modules include production, slaughter, and preparation. The exposure assessment part is the focus of further analyses by application of different sensitivity analysis methods in Chapters 4 to 10. Selection of the exposure assessment part for the sensitivity analyses in the *E. coli* food safety risk assessment model can be justified since that in the hazard and risk characterization parts, the dose-response relationship is estimated as the output of the model, based on the results of the exposure assessment part and available surveillance data exist on the annual number of illnesses due to infection with *E. coli O157:H7*. Hence, uncertainty about the results provided by the exposure assessment part can lead to uncertainty in the final output of the model, which is the dose-response equation for *E. coli O157:H7*.

### 3.2.1 Production Module

The production module estimates the prevalence of *E. coli*-infected cattle entering US slaughter plants. A determination of the quantitative association between the incoming status of cattle and the outgoing status of harvested meat is the main objective of this module. Estimation of the proportion of *E. coli*-infected cattle at slaughter begins with estimation the proportion of infected cattle on the farm.

The prevalence of the infected cattle entering the slaughter plants may be reduced through actions on the farm or feedlot. Mitigation strategies typically target herd-level risk factors for *E. coli* control. As an example, vaccination for this pathogen would likely be applied at the herd level. Culled breeding cattle and feedlot cattle are separately modeled, because there is evidence showing that there may be differences in *E. coli* prevalence between these two types of cattle.

The following key terms are used throughout this module (FSIS, 2001):

- Infected Cattle: refers to cattle whose intestinal tracts are colonized with the *E. coli O157:H7* organisms.
- Contaminated Cattle: refers to cattle whose hides, hair, or hooves have some *E. coli O157:H7* organism residing on them.
- Prevalence: the proportion of infected herds or individual cattle in a population.
- Herd Prevalence: the proportion of herds with one or more *E. coli*-infected cattle when the reference population is all herds of one type, for example, breeding herds.
- Apparent herd prevalence: the proportion of herds with one or more test-positive cattle detected among all herds sampled. Test-positive samples include both infected and contaminated cattle.
- True herd prevalence: is estimated by adjusting apparent herd prevalence observed in surveys with herd sensitivity.
- Herd sensitivity: is the proportion of infected herds that, when tested, are detected as *E. coli* positive. Herd sensitivity is dependent on the number of samples collected within herds and the detectable prevalence of infected animals in the infected herd.
- Within herd prevalence: is the proportion of infected cattle when the reference population is the cattle within a specific infected herd. This measurement is only applied to infected herds.

- True within herd prevalence: is estimated by adjusting apparent within-herd prevalence by test sensitivity.
- Test sensitivity: proportion of infected cattle, when tested, are detected as *E. coli*-positive using a particular diagnostic test.

The production module is comprised of three segments: (1) on-farm; (2) transportation; and (3) slaughter plant intake. The on-farm segment is comprised of four parts for estimating: (1) true herd prevalence; (2) true feedlot prevalence; (3) true within breeding herd prevalence; and (4) true within feedlot prevalence. Variability of true within herd or feedlot prevalence among all infected herds or feedlots and the seasonal variability of true within herd or feedlot prevalence are also estimated.

The four critical inputs to the production module are herd prevalence, within herd prevalence, feedlot prevalence, and within feedlot prevalence of *E. coli*. The production module simulates cattle entering the slaughter process via truckloads. Therefore, the prevalence of infection within truckloads is the module's output and the first input to the slaughter module. In Figure 3-2 the connectivity between different segments of the production module is depicted in a flowchart.

### 3.2.2 Slaughter Module

The slaughter module estimates the occurrence and extent of *E. coli* contamination as live cattle transition to carcasses, meat trim, and aggregates of meat trim in 60-pound trim boxes or 2000-pound combo bins destined for commercial ground beef production. Two types of slaughter plants are modeled: (1) those that handle culled breeding cattle (cow and bull); and (2) those that handle feedlot cattle (steer and heifer). The model only considers the commercial slaughter and processing of cattle. Prevalence distributions of *E. coli* in breeding and feedlot cattle, developed in the production module, serve as inputs to the slaughter module.

The prevalence distributions provide the number of infected cattle entering the slaughter plant. Breeder and feedlot cattle slaughtering operations are modeled separately, as are high and low prevalence seasons.

The following key terms are used throughout this module (FSIS, 2001):

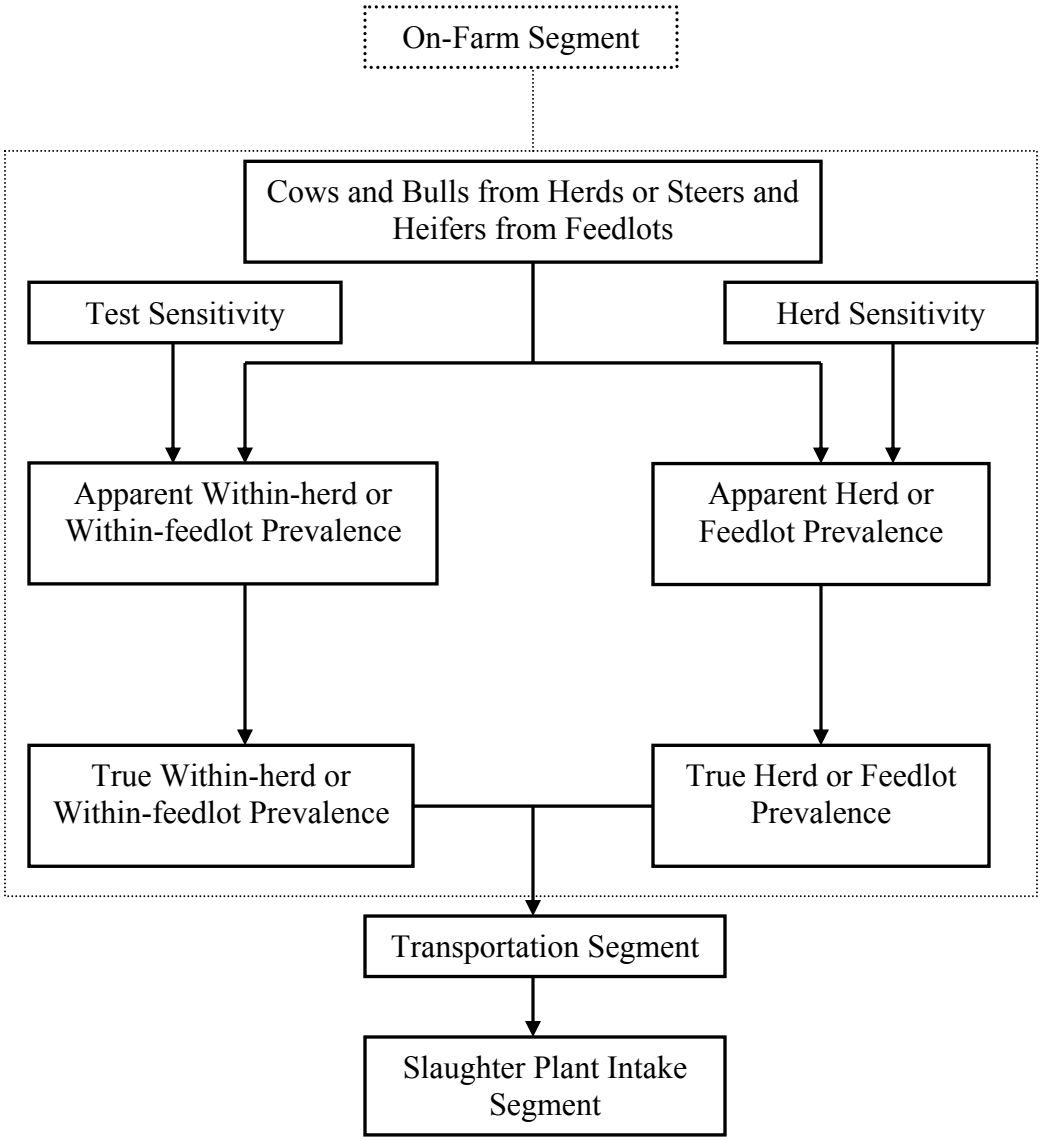



Figure 3-2. Schematic Diagram of the Production Module for the *E. coli O157:H7* Food Safety Risk Assessment Model.

- Carcass: refers to an animal that has been killed and had its hide removed.
- Contamination: is the presence of *E. coli* on carcass surface.
- Trim: is a by-product of processing carcasses to create cuts of meat when the carcasses originate from feedlot cattle. Trim consists of both muscle and fat.
- Combo bins: are containers that hold 2000 pounds of meat trim. Many cattle may contribute meat trim to a single combo bins.
- Boxes: of meat trim are similar to combo bins, but only contain 60 pounds of product.

- Lot**:** is defined as total number of cattle necessary to fill one combo bin. A single lot may comprise one or more truckloads of cattle

The slaughter module includes seven steps: (1) arrival of live cattle at the slaughter plant; (2) dehiding; (3) decontamination following dehiding; (4) evisceration; (5) final washing; (6) chilling; and (7) carcass fabrication. The module assumes that either contamination or decontamination can occur at each step of the process, with the prevalence and extent of contamination increasing if further contamination occurs and decreasing if decontamination occurs. The probability and extent of *E. coli* contamination or decontamination during slaughter are modeled as dependent on the status of the incoming animal, type of processing plant, type of equipment and procedures used, efficiency of decontamination procedures, and sanitation processes. In Figure 3-3 the connectivity between different parts of the slaughter module is depicted in a flowchart.

In step 1, cattle arrive at slaughter plants via truckloads with variability in the prevalence of infected cattle. In step 2, dehiding, there is the transition from live cattle to carcasses. This process creates the first opportunity of contamination of the carcass with *E. coli*. The number of *E. coli* organisms that initially contaminate a carcass depends on the level of infected cattle, the average concentration of the pathogen per contaminated area, and the total area of the carcass that is contaminated. In step 3, the number of *E. coli O157:H7* organisms on contaminated carcass surfaces can be reduced by the decontamination processes, including trimming, vacuuming, and washing of the carcass surface. Step 4, evisceration, is another opportunity for contamination to be introduced. Following final washing in step 5, the carcasses move to step 6, which is chilling. During the chilling process *E. coli* contamination may again increase or decrease. In step 7 the carcasses are fabricated. Because carcasses from breeding cattle produce less valuable whole muscle cuts than those from feedlot cattle, greater numbers of these deboned carcasses contribute to ground beef. The boneless meat trim from one animal is distributed based on fat content into multiple combo bins or boxes, where it is mixed with trims from other cattle.

Outputs from the slaughter module are distributions describing the frequency of *E. coli* in combo bins (and trim boxes) generated during high and low prevalence seasons for cow/bull and steer/heifer slaughter plants. These outputs are inputs to the preparation module. In Figure 3-3 the connectivity between different parts of the slaughter module is depicted in a flowchart.

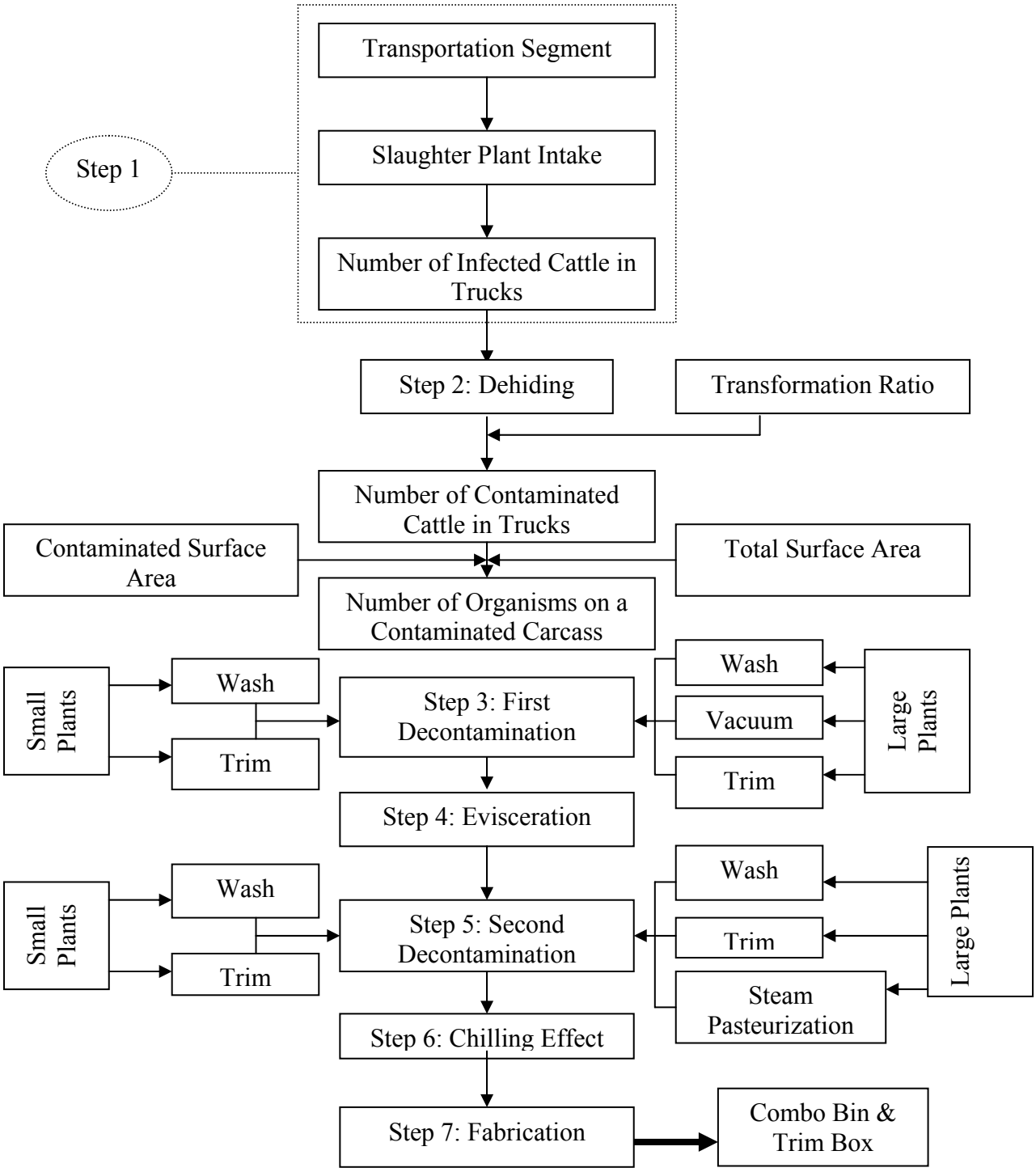



Figure 3-3. Schematic Diagram of the Slaughter Module for the *E. Coli O157:H7* Food Safety Risk Assessment Model.

### 3.2.3 Preparation Module

The preparation module estimates the occurrence and extent of *E. coli* contamination in consumed ground beef servings. This module also characterizes the consumption of ground beef servings by the age of consumer and the location of the meal.

The preparation module simulates the annual consumption of approximately 18 million ground beef servings. The model focuses on ground beef in the form of hamburger patties, and on ground beef as a formed major ingredient. Although cross-contamination could be a potential contributor for contamination of ground beef product, cross-contamination of ground beef products is not modeled. Cross contamination is the transfer of harmful microorganisms to food. It can occur in many ways, including contact from human hands, use of unsanitary equipment or work surfaces, storage or raw foods above ready-to-eat foods, or use of unsanitary cleaning cloths. In Figure 3-4 the connectivity between different parts of the preparation module is depicted in a flowchart.

The following key terms are used throughout this module (FSIS, 2001):

- Servings: the amount of ground beef consumed per eating occasion. It varies by the age of consumer and the location where the meal is consumed (e.g., at home versus away from home).
- Exposure: amount of contamination that is consumed in a serving.
- Home: is used when servings are prepared and served in a home environment.
- Away from home: is used when servings are prepared and served in an institutional environment. This is often referred to as "HRI" (hotels, restaurants, and institutions).
- Transportation: refers to non-refrigerated transport of product from a retail to wholesale establishment.
- Retail: refers to establishments, such as grocery stores or butcher shops, that sell ground beef for home consumption.
- Wholesale: refers to establishments that serve as distributors to HRI for away from home consumptions.

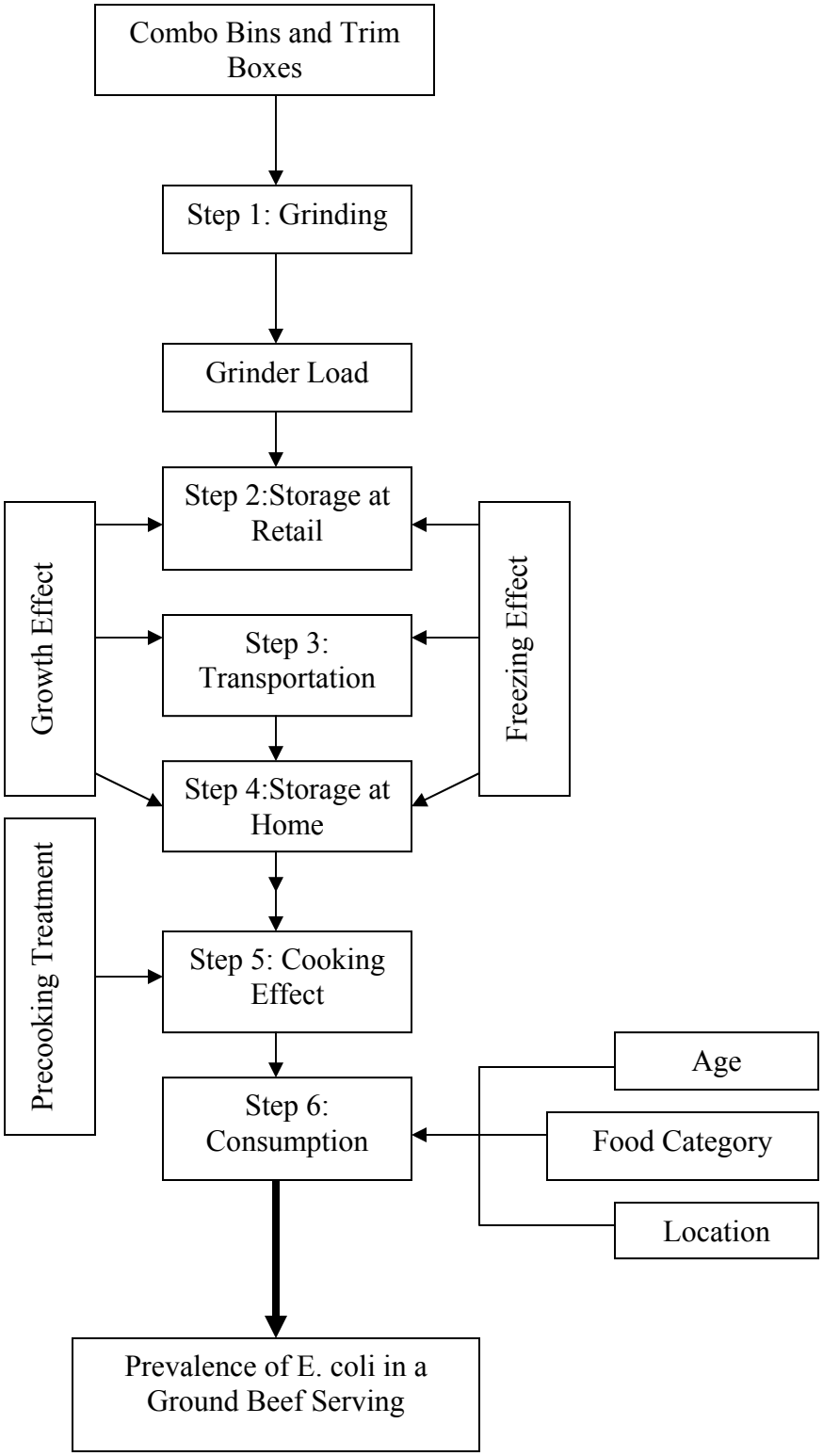



Figure 3-4. Schematic Diagram of the Preparation Module for the *E. Coli O157:H7* Food Safety Risk Assessment Model.

The preparation module consists of six primary steps. Five of these steps explicitly model growth, decline, or dispersion of *E. coli O157:H7* contamination: (1) grinding; (2) storage at retail; (3) transportation; (4) storage at home or away from home; and (5) cooking. Step 6 models the amount of ground beef consumed, which varies depending on the age of the consumer and the eating location. In step 1, multiple combo bins or boxes are combined and mixed to produce finished ground beef with a specific fat content. Although the extent of *E. coli* contamination does not increase during the grinding process because of temperature controls, contamination from single combo bins or boxes can be dispersed during grinding to contaminate many individual ground beef servings. In step 2, storage conditions at retail or wholesale provide an opportunity for *E. coli O157:H7* levels to increase as a result of increased time and temperature or decrease as a result of freezing ground beef. High storage time or storage temperature at retail leads to increase in the number of *E. coli O157:H7* organisms. Step 3 models the effects of time and temperature during transportation on the level of *E. coli O157:H7* after the ground beef is purchased. Step 4 models the storage of ground beef at freezer and refrigerator prior to its preparation and consumption. In step 5, the effect of cooking on the number of *E. coli O157:H7* organisms is evaluated. Step 6 models the consumption of contaminated ground beef servings.

An intermediate output of the preparation module is the distribution of *E. coli* densities in grinder loads of ground beef made from 2000-pound combo bins. Another intermediate output of the preparation module is the distribution of *E. coli* densities in grinder loads of ground beef made from 60-pound trim boxes. The primary outputs from the preparation module are distributions describing the prevalence of *E. coli O157:H7* in ground beef servings generated during low and high prevalence seasons (winter and summer, respectively).

### 3.2.4 Limitations

The *E. coli* food safety risk assessment model was not originally developed for the purpose of facilitating sensitivity analysis. The objective of this section is to identify critical needs for sensitivity analysis and to determine the limitations of the existing model with respect to these needs. Based upon these limitations, specific requirements are identified for modifying the existing model in order to facilitate sensitivity analysis. The modifications are documented in Section 3.4.

One of the most important goals of sensitivity analysis, as described at the NCSU/USDA Workshop on Sensitivity Analysis, is to perform global sensitivity analysis on output variables of

direct relevance to a decision. However, the *E. coli* risk assessment model is implemented with separate modules that make it impractical to perform global sensitivity analysis upon the entire model.

A second desirable goal of sensitivity analysis is to distinguish between variability and uncertainty. Many of the inputs to the *E. coli* model can be conceptualized as representing variability only, uncertainty only, or both variability and uncertainty. However, the manner in which the probabilistic analysis was implemented for the *E. coli* model makes these distinctions difficult in the context of a single simulation. Rather, in order to distinguish between variability and uncertainty with the existing model, it is necessary to run the model for separate case studies of variability only, uncertainty only, variability for different uncertainty realizations, or co-mingling of both variability and uncertainty in a single probabilistic simulation.

A third limitation is that data for many of the intermediate variables are binned. Therefore, it is not possible to trace the value of a model output to a specific combination of input values. Thus, because the model is structured to bin intermediate results, it is not possible to have a one-for-one correspondence between the value of a model output and the values assigned to model inputs, which poses a challenge for performing sensitivity analysis. In summary, the three key limitations of the *E. coli* model with respect to sensitivity analysis include: (1) modularity of the model based upon division into modules; (2) challenges in distinguishing between variability and uncertainty; and (3) coding limitations pertaining to binning of intermediate inputs. Each of these limitations is described in more detail.

#### 3.2.4.1 Modularity

The *E. coli* risk assessment model is divided into modules, as described in Sections 3.1.1 through 3.1.3. Outputs of one module serve as inputs to the next. In combination with the fact that many of the intermediate values of variables are binned, the implication of both modularity and binning of variables is that there is a lack of one-for-one correspondence between the value of a desirable risk assessment model output, such as contamination in ground beef servings, and the values of inputs to the various modules that influence the output. The possibility of modifying the model to avoid binning of intermediate variables and to allow for direct communication of data from one module to the next was considered and explored.

The original version of the *E. coli* food safety risk assessment model was implemented in Microsoft Excel using inter-cellular functioning by implementing equations inside worksheet

cells prepared for each module. Furthermore, the @Risk software was used to define probability distributions for inputs to the model. The use of equations inside worksheet cells, as opposed to the use of a stand-alone programming language, permits execution of the code in the spreadsheet environment; however, spreadsheet-based models are difficult to modify compared to programming language-based models. In order to perform sensitivity analysis, a data set should be formed consisting of the desired candidate inputs and the outputs of interest. Thus, it is desirable to easily access and store paired data values for the inputs and outputs in order to facilitate sensitivity analysis. In the existing model, data are not routinely stored for this purpose. Therefore, in order to save the data needed for sensitivity analysis, the code in different modules and parts of the *E. coli* model is modified to save such data.

#### 3.2.4.2 Challenges in Distinguishing Between Variability and Uncertainty

In risk assessment, it is often useful to distinguish between variability in exposure and risk versus uncertainty regarding knowledge of the true value of a quantity or distribution. An accepted method for distinguishing between variability and uncertainty in human health risk assessment is to use two-dimensional probabilistic simulation (e.g., Frey and Rhodes, 1996). This approach requires that each input to the model be appropriately simulated to represent either variability and/or uncertainty. In the following this issue is clarified with an example.

In a two-dimensional probabilistic simulation each input variable has a two-dimensional matrix, containing the generated values for the input variable for both variability and uncertainty simulations. In Figure 3-5, this two-dimensional matrix for input $X$ is depicted. In this example, variable $X$ has both variability and uncertainty.

The number of iterations for the variability simulation is $n$, and the number of iterations for the uncertainty simulation is $m$. Hence, an $n * m$ matrix is generated for variable $X$, with columns representing variability, and rows representing uncertainty. For instance, in Figure 3-5, the first column represents the first uncertainty iteration, and $n$ variability iterations. In this column, the uncertain part of the input $X$ remains constant.

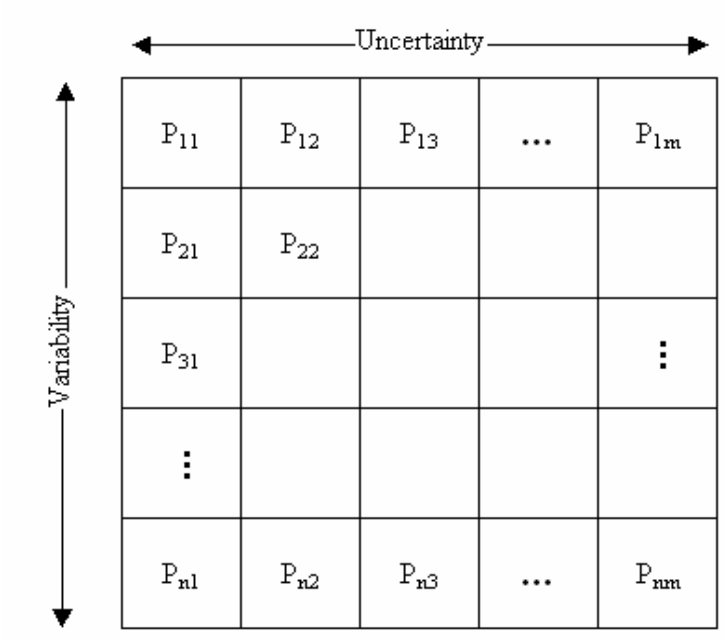



Figure 3-5. Matrix of Generated Values for Variable *X* in a Two-Dimensional Simulation.

As an example, Figure 3-6 depicts the variability distribution of the cooking temperature as an input in the preparation module for the $j^{th}$ uncertainty iteration. During the variability simulation cooking temperatures are selected from the distribution based on random samples. Hence, for *n* variability iterations there are *n* values for the cooking temperature. These values are placed in different columns of the matrix in Figure 3-5, based on the number of the uncertainty iteration (e.g., j =1 to m). Figure 3-7 depicts the effect of the uncertainty in the cooking temperature distributions. At a specific variability percentile, different cooking temperatures can be generated, because of the uncertainty in the cooking temperature distributions. Hence, for *m* uncertainty iterations there are *m* values for the cooking temperature at each variability percentile. These values are placed in different rows of the matrix in Figure 3-5, indicating that each row represents a specific variability percentile.

In order to form the above matrix with columns representing the variability and rows representing the uncertainty, the same set of random numbers for the variability simulation should be used in different uncertainty iterations. Using the same set of random numbers facilitates stratifying between variability and uncertainty in the matrix, since for example the variation in each row can be attributed to only the uncertainty in specific percentile, and not the contribution of both uncertainty and variability.

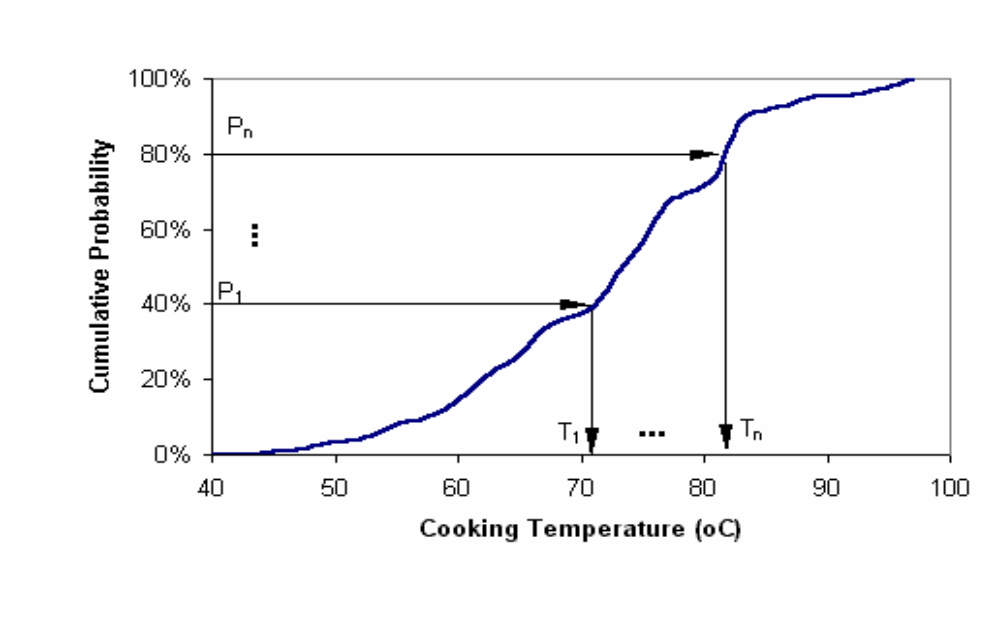



Figure 3-6. Cumulative Variability Distribution for the $j^{th}$ Uncertainty Iteration.

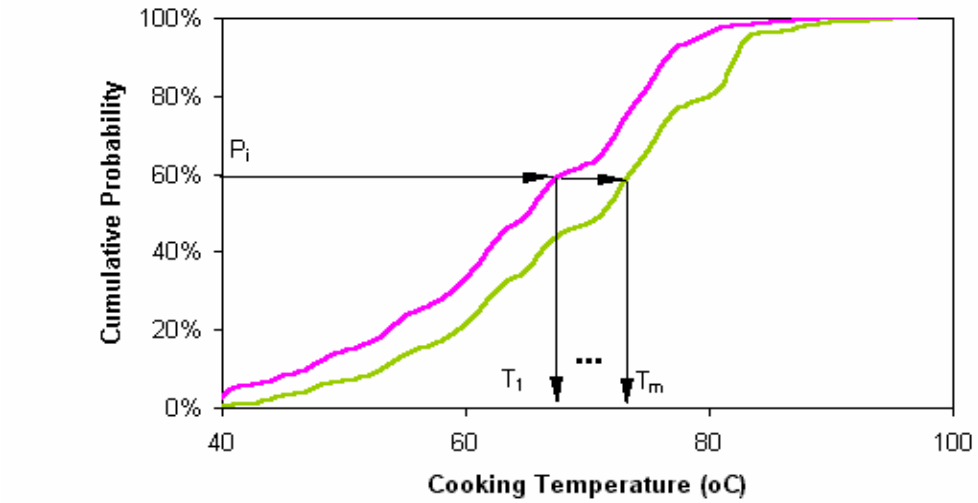



Figure 3-7. Different Values of the Cooking Temperature at Specific Variability Percentile, Representing the Uncertainty about the Input.

However, the *E. coli* model is structured in a way that makes it difficult to fully distinguish between uncertainty and the variability. In the *E. coli* model, the random numbers used during the variability iterations are not stored. Hence, for different uncertainty iterations, different sets of random numbers are used for the variability simulation. Thus, variation of the

numbers in each row of the matrix cannot be attributed only to the influence of uncertainty in the input distribution, because they also are influenced by a random component within the variability simulation. Hence, it is impractical to perform analysis of uncertainty only in the modules and parts of the original *E. coli* model that have two-dimensional simulation such as the slaughter module and the growth estimation part.

The *E. coli* model can be configured to run case studies in a one-dimensional framework to simulate only variability in each input or only uncertainty of each input for modules and parts that have one-dimensional simulation, such as the production module, cooking effect or serving contamination parts. Furthermore, the model can be run for multiple realizations of variability based upon different estimates of uncertainty. Thus, four types of case studies are included in later chapters: (1) variability only; (2) uncertainty only; (3) variability and uncertainty in separate dimensions, with a focus on how uncertainty impacts the realizations of variability; and (4) variability and uncertainty combined into one-dimension, representing a randomly selected individual.

#### 3.2.4.3 Coding Limitations Pertaining to Binning of Intermediate Inputs

Because many intermediate inputs in the *E. coli* risk assessment model are binned, it is not possible to trace the influence of specific values of model inputs to corresponding values of model outputs. For example, the contaminant concentration in combo bins is simulated as a continuous variable, but subsequently is binned into increments of 0.5 logs ranging from 0 to 8 logs. 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 considered as 0.5 logs, and if it is 1.01 logs, it is considered as 1.5 logs. In this way, when the combo contamination is used as input in the next modules, and a value is selected from its binned distribution, there is no way to identify the original value of the combo contamination before the binning process. This issue makes back tracking, which is essential for developing the dataset for the sensitivity analysis, almost impossible. In order to eliminate the bining approach, it would be necessary to substantially change the model structure. Such a change was beyond the scope of this work. Therefore, global sensitivity analysis could not be performed in the *E. coli* model. Instead, different case scenarios for sensitivity analysis are focused on different modules and parts of the *E. coli* model individually. In Chapter 10 results of the local sensitivity analyses in different modules and parts

of the *E. coli* model are evaluated in order to come up with general conclusions regarding the relative significance of different parts of the model.

In the next section, key scenarios that were selected as the basis for case studies with the *E. coli* model are identified. The focus of the selected case studies, combined with the limitations described in this section, were used as a basis for prioritizing activities to modify the *E. coli* model as appropriate in order to facilitate sensitivity analysis. The modifications are documented in Section 3.5.

## 3.3 Case Scenarios

Any analysis calls for resources both in terms of time and space. Hence, it is important to identify the highest priority scenarios needed for evaluation. For example, one season may provide more ideal conditions for pathogen organism dispersion and spread of the disease than others. A sub-population may be more susceptible to the adverse effects caused by consumption of the contaminated food. The methodology for sensitivity analysis is not dependent on these factors. Thus, for purpose of demonstrating methods, it is not necessary to consider all possible scenarios. Furthermore, in order to have meaningful outcomes from the sensitivity analysis that can be used by food scientists, it is useful to define specific case scenarios relevant to the model scope that are of policy interest. Therefore, the objective of this section is to define specific scenarios that are the focus of sensitivity analysis case studies. The scenarios are defined for each of the three exposure assessment modules, including production, slaughter, and preparation.

### 3.3.1 Production Module

In this section the case scenario in the production module is explained. The explanation of the case scenario includes the identification of cattle categories, and seasons that are considered in the analysis. The characteristic of the simulation, regarding the incorporation of variability, uncertainty, or both is specified, and the number of iterations in the simulation is introduced. In addition, a few questions are raised at the end of the section. These questions are addressed later based on the sensitivity analysis methods that are applied to this module in Chapters 4 to 10.

In the production module both feedlots cattle (e.g., steers and heifers) and breeding herds cattle (e.g., cow and bulls) are considered in the analysis. Regarding the temporal dimension of the analysis, both high and low prevalence seasons are considered in the analysis. In the production module there is a one-dimensional uncertainty simulation. The number of iterations

in the simulation equals 65,000. This number is selected considering the maximum possible number of rows in an EXCEL worksheet, since the generated data during the simulation are stored in an EXCEL worksheet.

Key questions are raised here for the production module, and these questions are answered in Section 11.2.1 based on results of different sensitivity analyses in Chapters 4 to 10.

Question 1: What is the ranking of the input variables regarding their influence on the output of interest?

Question 2: Is there any study effect in estimation of the response?

Question 3: Is there any seasonality effect for estimation of average within feedlot or breeding herd prevalence?

Question 4: Which of the testing methods provide higher accuracy?

Regarding Question 2, in different parts of the production module, such as feedlot or breeding herd prevalence, and within feedlots or breeding herd prevalence, several studies provide information about the population variation for the infection prevalence. These studies have specific characteristics regarding the number of samples, number of positive cases and specific testing methods, implemented in the study. In the original *E. coli* food safety risk assessment model, the final outputs in the production module incorporate the effect of different studies in estimation of feedlot or breeding herd prevalence and within feedlots or breeding herds prevalence. Since some of the studies may have greater credibility because of the larger sample size and better testing methods with higher accuracy, the effect of a choice among the studies as information sources on the final outputs of the module was evaluated. The results from this analysis provide insight regarding the impact of selection of different studies as a basis for estimating the infection prevalence. The information regarding different studies implemented in the production module is given in Tables 3-1 to 3-4.

### 3.3.2 Slaughter Module

In this section, the case scenario in the slaughter module is explained. For explanation of the case scenario, cattle categories that are considered in the analyses are identified. In addition, the season selected for the analyses and the reason for this selection is presented. The characteristic of the simulation regarding the incorporation of the variability, uncertainty, or both is specified, and the number of iterations in the simulation is introduced. In addition, a few

Table 3-1. Different Studies Used for Estimation of the Feedlot Prevalence

| Study | Dargatz Hancock 1997 | Hancock 1998 | Smith 1999 | Elder 2000 |
|---|---|---|---|---|
| Number of Positive Cattle in Positive Herds | 210 | 38 | 707 | 91 |
| Number Tested in Positive Herds | 7560 | 1046 | 3054 | 254 |
| Samples per Feedlot | 120 | 174 | 611 | 12 |
| Number of Feedlots Tested | 100 | 6 | 5 | 29 |
| Positive Feedlots | 63 | 6 | 5 | 21 |

Source: Table 3-6, *E. coli* food safety risk assessment model report, FSIS 2001

Table 3-2. Different Studies Used for Estimation of the Within Feedlot Prevalence

| Study | Dargatz Hancock 1997 | Hancock 1999 | Hancock 1998 | Smith 1999 | Elder 2000 |
|---|---|---|---|---|---|
| Total Month with Available Data | 3 | 5 | 5 | 4 | 2 |
| Testing Method | 0.1g, SMACct | 0.1g, SMACct | 0.1g, SMACct | 10g, IMS | 10g, IMS |
| Study Weight | 2520 | 48 | 209 | 764 | 127 |

Source: Table 3-6, *E. coli* food safety risk assessment model report, FSIS 2001

Table 3-3. Different Studies Used for Estimation of the Breeding Herd Prevalence

| Study | Garber 1998 | Sargeant 2000 | Hancock FDA 2001 | Hancock 1997a | Hancock 1998 | Lagreid 1998 | Hancock 1997b |
|---|---|---|---|---|---|---|---|
| Number of Positive Cattle in Positive Herds | 51 | 29 | 38 | 179 | 25 | 61 | 91 |
| Number Tested in Positive Herds | 1268 | 2348 | 5709 | 9720 | 1097 | 758 | 7121 |
| Samples per Herd | 58 | 235 | 317 | 360 | 183 | 60 | 791 |
| Number of Herds Tested | 91 | 10 | 20 | 36 | 6 | 15 | 13 |
| Positive Herds | 22 | 10 | 18 | 27 | 6 | 13 | 9 |

Source: Table 3-2, *E. coli* food safety risk assessment model report, FSIS 2001

Table 3-4. Different Studies, Used for Estimation of the Within Breeding Herd Prevalence

| Study | Garber 1998 | Besser 1997 | Rice 1997 | Hancock 1994 | Sargeant 2000 | Hancock FDA 2001 |
|---|---|---|---|---|---|---|
| Total Month with Available Data | 5 | 12 | 6 | 10 | 12 | 8 |
| Testing Method | 1g SMACct TSB | 0.1g SMACct | 0.1g SMACct | 0.1g SMAC | 10g IMS | 0.1g SMACct |
| Study Weight | 254 | 173 | 13 | 46 | 196 | 714 |

Source: Table 3-2, *E. coli* food safety risk assessment model report, FSIS 2001

questions are raised at the end of the section. These questions are addressed later based on the sensitivity analysis methods that are applied on this module in Chapters 4 to 10.

The output of interest in the slaughter module is the contamination in combo bins and trim boxes. Cattle harvested from feedlots have higher probability of infection than cattle from breeding herds, and regardless of cattle type, more highly infected cattle enter the slaughter plants during the high prevalence season (FSIS, 2001). Therefore, feedlot cattle in the high prevalence season are selected as the focus of the case study.

In the slaughter module there is a two dimensional simulation, incorporating both variability and uncertainty. In order to have profound insight regarding the individual variability and uncertainty effect on the output of the module, three different analyses are performed: (1) variability analysis for a mean uncertainty effect; (2) variability analysis for several uncertainty iterations; and (3) mixed analysis.

In the first analysis, in order to evaluate the sole effect of the variability in inputs of the slaughter module on the output of interest, the uncertainty in the inputs are fixed at point estimates (e.g., mean value of the uncertainty distribution). The number of iterations in the variability analysis simulation equals to 65,000, based on the upper limit of what can be simulated in EXCEL.

In the second type of analysis, the objective is to distinguish between variability and uncertainty. The focus of the sensitivity analysis is to identify the key inputs for each realization

of variability. Specifically, the model is executed 650 times for each estimate of variability, and this is repeated 100 times for different estimates of uncertainty. Thus, sensitivity analysis is applied 100 times to identify key inputs. To the extent that the sensitivity analyses yield similar results regarding 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. Such a result would imply that an analyst or decision maker could make a robust determination as to what model inputs are contributing the most to variability in exposure and, hence, risk and as to what CCP's could be employed to reduce the high end exposures. In contrast, if the ranking of key inputs changes substantially from one simulation of uncertainty to another, then the identification of key inputs would be ambiguous because of uncertainty. The latter result implies that there could be a benefit to targeting data collection or research so as to reduce uncertainty. The sample sizes for variability and uncertainty were selected based upon the constraint that only 65,000 simulations can be carried out in the Excel-based model. A sample size of 100 for the uncertainty dimension is adequate to capture a wide range of variation in uncertain model inputs. The uncertainty sample size was kept smaller than the variability sample size because the uncertainty sample size also determines the number of iterations, and hence the computational burden, for the sensitivity analysis methods. Because only a few of the selected sensitivity analysis methods were readily amenable to automation, the two-dimensional framework for dealing with variability distinct from uncertainty was applied only with the regression analysis and ANOVA techniques.

The third type of analysis involves co-mingling of both variability and uncertainty into a one-dimensional probabilistic simulation. This type of analysis can be used to address issues pertaining to a randomly selected individual. Many key risk characterization questions focus either upon the average risk to a population or the high end risk among individuals. Such questions cannot be answered based upon analysis of a randomly selected individual. However, as an aid to sensitivity analysis, analysis of a randomly selected individual may have some advantages. In particular, by co-mingling both variability and uncertainty into a single dimension, one typically would obtain wider ranges of values for model inputs than if only uncertainty or only variability had been characterized. Thus, in a single probabilistic simulation it is possible to exercise as fully as possible the range of model inputs and the corresponding impact on model outputs. This type of analysis could be used, for example, as a screening tool to

perhaps focus subsequent two-dimensional analyses upon those inputs that appear to matter the most.

Key questions are raised here for the slaughter module, and these questions are answered in Section 11.2.2 based on results of different sensitivity analyses in Chapter 4 to 10.

Question 1: What is the ranking of inputs regarding their influence on the output of interest?

Question 2: How robust is the identification of key inputs for situations in which variability and uncertainty can be distinguished?

Question 3: Which step in the slaughter module could produce high contamination levels in combo bins?

Question 4: How can the decontamination steps mitigate the number of *E. coli* organisms in combo bins?

### 3.3.3 Preparation Module

In this section the case scenarios in the preparation module are explained. The preparation module consists of three main parts: (1) growth estimation; (2) cooking effect; and (3) serving contamination. Different case scenarios are presented for each part in the following text. Explanation of case scenarios includes identification of stages considered in each part, ground beef consumption types, sub-population age groups, and the temporal dimension of the analysis. The characteristic of the simulation regarding the incorporation of variability, uncertainty, or both is specified. In addition, a few questions are raised at the end of the section. These questions are addressed later based on the sensitivity analysis methods that are applied to this module in Chapters 4 to 10.

In the growth estimation part, three stages are considered: (1) retail; (2) transportation; and (3) home. There is no temporal dimension in the growth estimation part, indicating that no difference was considered in the growth of *E. coli O157:H7* organisms in different seasons. The growth estimation part has a two-dimensional simulation incorporating both variability and uncertainty in the analysis. Three different probabilistic analyses are performed in the growth estimation part similar to the approach for the production module: (1) variability analysis based upon mean point estimates for uncertainty; (2) variability analysis for alternative realizations of uncertainty; and (3) a one-dimensional analysis in which variability and uncertainty are combined. The explanation of each type of analysis was presented in Section 3.2.2.

In the cooking effect part, nine precooking treatments are considered in the analysis in order to identify the effect of precooking treatments on the log reduction in the number of *E. coli O157:H7* organisms. In addition, two cooking locations (i.e., home and away) are considered to identify whether there is any difference in the cooking effect when ground beef servings are cooked at home or away from home. The cooking effect part has a one-dimensional probabilistic simulation incorporating only variability. Hence, 65000 variability iterations were performed in this part.

The case study scenario in the serving contamination part includes consideration of three ground beef consumption types: (1) raw ground beef; (2) hamburger patties; and (3) meatballs. These ground beef consumption types are considered in order to compare the effect of different consumption types with respect to the serving contamination. In addition, four age groups (i.e., 0-5, 6-24, 25-64, 65+) are considered in the analysis. Children less than 5 years old and elderly people are considered as susceptible sub-populations to *E. coli O157:H7* based upon surveillance data (FSIS, 2001), but all four sub-populations are incorporated in the simulation in order to identify the effect of consumer age on the serving contamination. Eating at home and away from home are considered as two alternative eating locations, in order to identify any effect of eating place on the serving contamination. Serving contamination part has a one-dimensional probabilistic simulation incorporating only variability. Hence, 65000 variability iterations are performed in this part.

Key questions are raised here for the preparation module, and these questions are answered in Section 11.2.3 based on results of different sensitivity analyses in Chapters 4 to 10.

Question 1: What is the ranking of the input variables regarding their influence on the output of interest in different parts of the module?

Question 2: How robust is the identification of key inputs for situations in which variability and uncertainty can be distinguished?

Question 3: What is the effect of precooking treatments on the log reduction due to cooking?

Question 4: How does the contamination level differ for different age groups?

Question 5: What is the effect of eating location on the possible contamination of a ground beef serving?

Question 6: Does the eating place affect the contamination in different ground beef consumption types?

## 3.4 Modifications to Enable Sensitivity Analysis

Based on the discussion of the model's limitations in Section 3.1.4, the original *E. coli* code was modified in order to prepare the *E. coli* food safety risk assessment model for performing deferent sensitivity analysis methods. The modifications were done so as not to change the original model structure. The modifications are classified into three parts, including modifications in the: (1) production module; (2) slaughter module; and (3) preparation module. Each of the modifications is discussed in the following subsections.

### 3.4.1 Production Module

Based on the case scenario for the production module as described in Section 3.2.1, the intent was to identify the study effect and the seasonality influence on the outputs of interest. Modifications in the production module are classified into two parts: (1) modifications in the feedlot or breeding herd prevalence estimation part; and (2) modifications in the within feedlots or breeding herds prevalence estimation part.

For the feedlot or breeding herd prevalence estimation part, there is no seasonality effect, since the temporal variation of these outputs was not considered in the *E. coli* model. Hence, the original model is modified in a way that the study effect can be evaluated. The original *E. coli* model considered equal weights for different studies used for estimation of the feedlot or breeding herd prevalence. In order to consider the study as a random variable in the modified *E. coli* code, a discrete distribution with equal weights was defined for the study effect. Figure 3-8 depicts the modified algorithm for the estimation of feedlot or breeding herd prevalence.

Based on the modified algorithm, in step 1 of each iteration, a study is picked from the discrete distribution defined for the study effect. In step 2, the herd sensitivity and the apparent feedlot or breeding herd prevalence is estimated based on the data available from the selected study. In the original *E. coli* model a distribution is estimated for the output, and then a number randomly is picked from that distribution as the value of feedlot or breeding herd prevalence for that iteration. In order to eliminate this random characteristic of the output, in step 3 the median value of the estimated output distribution is considered as the output of interest.

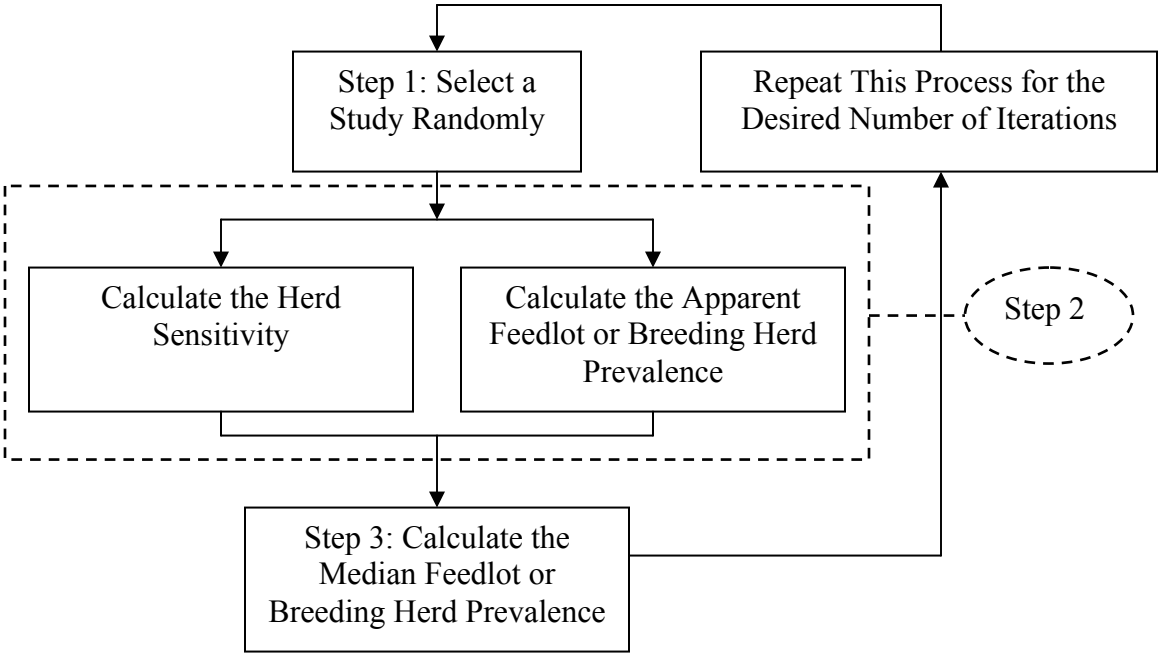



Figure 3-8. Schematic Diagram of Modified Algorithm for the Feedlot or Breeding Herd Prevalence Estimation.

For the within feedlot or breeding herd estimation part, both the seasonality effect and the study effect are considered. The original code averages over different studies and estimates the within feedlot or breeding herd prevalence for high and low prevalence season separately. In the original code the final output in this part is the average between different studies considering the number of months with available data as the weight of the study. In the modified version of the *E. coli* code, a discrete distribution with unequal weights based on the number of months with available data is defined for the study effect. Figures 3-9 and 3-10 depict the distributions defined for the study as a random input in estimation of average within feedlot and breeding herd prevalence. To account for the seasonality effect in the within feedlot or breeding herd prevalence, a discrete distribution with equal weights was defined. In this way, season participates in the simulation as a random variable. Figure 3-11 depicts modified algorithm for the estimation of average within feedlot or breeding herd prevalence.

Based on the modified algorithm, in step 1 of each iteration a study is picked from the discrete distribution defined for the study effect. In step 2, the season is selected from its distribution. In step 3, based on the study and the season that are already selected, the test sensitivity and the apparent within feedlot or breeding herd prevalence are calculated. In step 4,

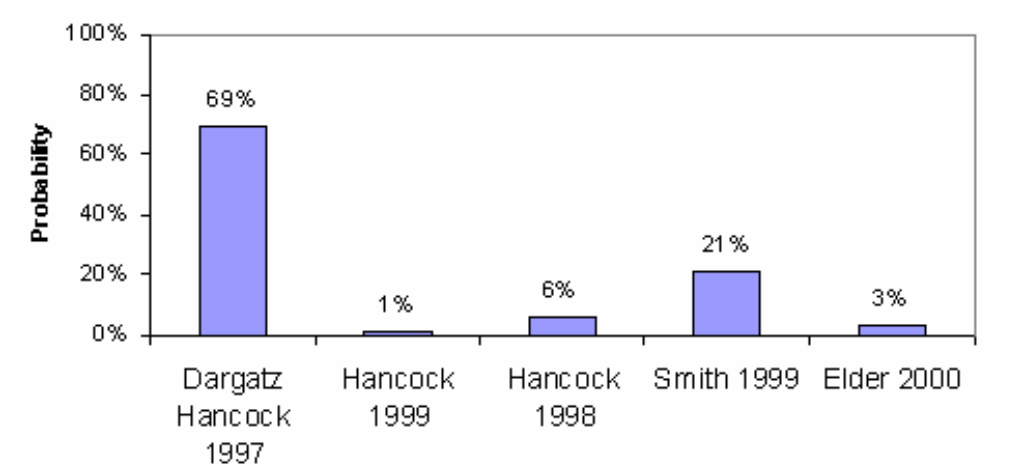



Figure 3-9. Discrete Distribution for the Study Effect in Within Feedlot Prevalence.

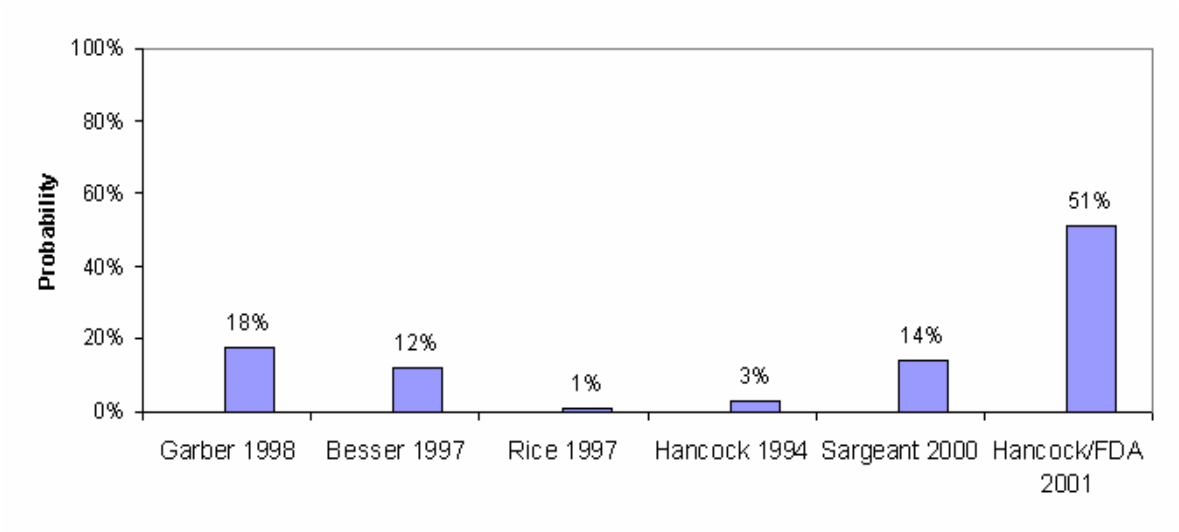



Figure 3-10. Discrete Distribution for the Study Effect in Within Breeding Herd Prevalence.

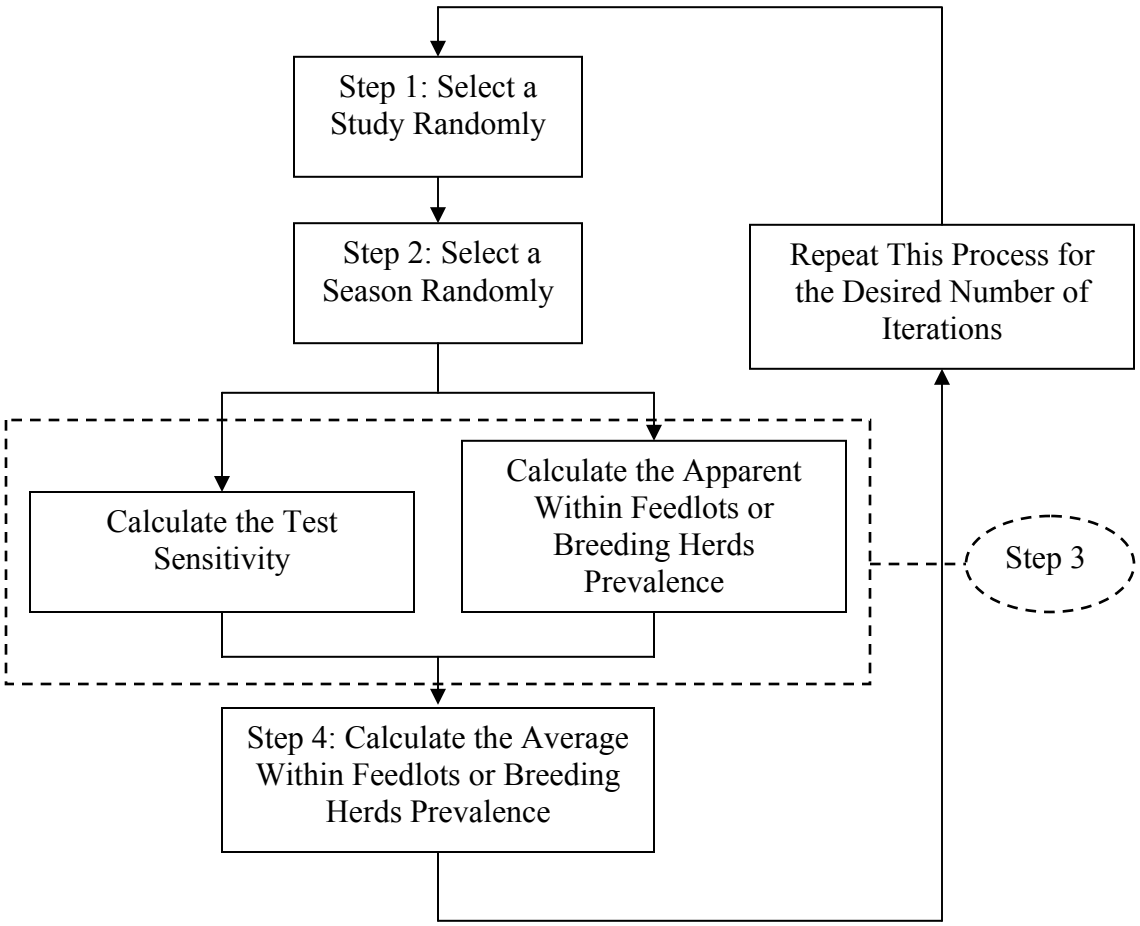



Figure 3-11. Schematic Diagram of Modified Algorithm for the Average Within Feedlot or Breeding Herd Prevalence Estimation.

the average within feedlot or breeding herd prevalence is estimated as the output of interest in this part.

### 3.4.2 Slaughter Module

In this section the modifications performed in the slaughter module are explained. In order to answer to the questions raised in Section 3.3.2 for the case scenario in the slaughter module, a dataset including the generated values of the relevant inputs in this module and the output of interest was prepared. That dataset has one column corresponding to each input, and a column for the output of interest. The number of rows in this dataset equals the number of iterations in a simulation. The formation of such a dataset is essential for performing any sensitivity analysis method. Sensitivity analysis methods typically require that for each input, one value is generated in an iteration and that there is an output corresponding to input variables

in an iteration. In the slaughter module of the original *E. coli* model it was not possible to form this type of dataset. In order to illustrate this issue, an example regarding the modeling approach in the slaughter module of the original *E. coli* model is presented here. At the end of this section, the modifications in the slaughter module in order to generate the required dataset are presented.

In the slaughter module contamination in combo bins and trim boxes is estimated, considering that combo bins and trim boxes are filled with meat trims coming from cattle that are slaughtered in the slaughter plant, dehided, eviscerated, and finally fabricated in different parts of the plant. The aggregation issue in filling combo bins and trim boxes with meat trims and the difficulty in tracking down the contaminated meat trims into different combo bins and trim boxes cause problem in generating the explained dataset. As an example, a combo bin may consist of 52 cattle. From these cattle, 42 may have no contamination, 5 are contaminated only at dehiding, 2 are contaminated only due to evisceration, and the rest are contaminated at both steps, dehiding and evisceration. Hence, in this combo bin *E. coli* organisms come from three pathways. During estimation of the number of organisms in this combo bin, each contamination pathway is taken into account separately and then the final contamination of this combo bin includes the organisms from all three pathways. There are inputs affecting the combo bin contamination that have to be calculated for each contamination pathway separately, such as the number of organisms on a contaminated carcass, Trim/Vacuum/Washing efficiency, organisms added due to evisceration, washing efficiency, and contaminated $cm^2$ of meat trims in a combo. Figure 3-14 depicts a schematic flow of the process of filling a combo bin with meat trims, based upon different sources with different pathways of contamination.

Based on the schematic diagram in Figure 3-12, contamination from 3 pathways is aggregated and eventually there is only one value of contamination for the combo bin. In each pathway, there is another source of aggregation of *E. coli O157:H7* organisms. For example, in the first pathway, there are five cattle contaminated during the dehiding process. For each cattle, the amount of *E. coli O157:H7* organisms that each animal contributes to the total contamination of the combo bin is calculated separately, and then sum of the contaminations over different animals is considered as the aggregate pathway contamination. Thus, although there is only one value of the combo contamination as the output in this iteration, there are multiple values of inputs associated with each animal and with the aggregate effect of multiple animals on each pathway. For instance, in this example, all input variables affecting the combo contamination are

calculated five times in pathway one, three times in pathway two, and two times in pathway three. A question can be raised as to which of these input values should be considered as the appropriate one to use for sensitivity analysis corresponding to the output in the iteration, since ideally there should be only one input value associated with each output.

In order to solve this problem, the original *E. coli* code in the slaughter module was modified in a way that a new individual animal is introduced, representing the characteristics of all contaminated cattle contributing to the combo bin contamination. Figure 3-13 depicts the individual cattle representing all the contaminated cattle in different pathways of the presented example. This representative animal contributes the total number of *E. coli O157:H7* organisms to the combo bin. Hence, the equations used for calculation of the number of *E. coli O157:H7* organisms were modified in order to take into account that the data regarding representative animal are needed instead of data for individual contaminated cattle. In these equations, n indicates the number of contaminated cattle contributing in the combo bin (e.g., 10 cattle in presented example):

$$\text{Organisms on Contaminated Carcass} = \sum_{i=1}^{n} (OC)_i \tag{3-1}$$

$$\text{Trim/Vacuum/Wash Efficiency} = \frac{\sum_{i=1}^{n} [(OC)_i * (TVW)_i]}{\sum_{i=1}^{n} (OC)_i} \tag{3-2}$$

$$\text{Evisceration Organisms Added} = \sum_{i=1}^{n} (EOA)_i \tag{3-3}$$

$$\text{Wash Percent Reduction} = \frac{\sum_{i=1}^{n} [(OC)_i * (WR)_i]}{\sum_{i=1}^{n} (OC)_i} \tag{3-4}$$

$$\text{Contaminated Cm}^2 \text{ in a Combo} = \sum_{i=1}^{n} (CC)_i \tag{3-5}$$

$$\text{Combo Contamination} = \sum_{i=1}^{n} (CCM)_i \tag{3-6}$$

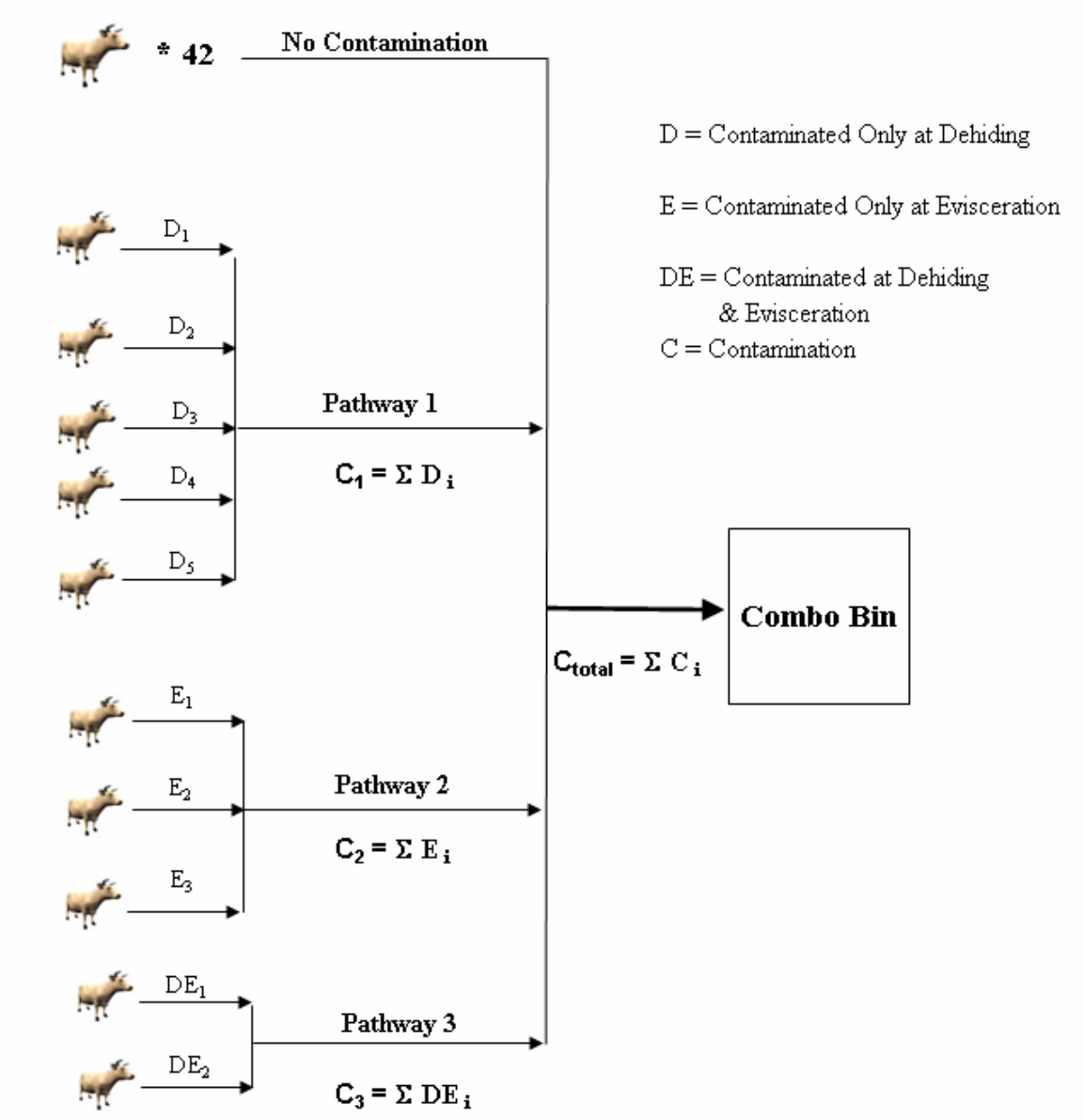



Figure 3-12. Schematic Flow of the Process of Filling a Combo Bin with Meat Trims Coming From Different Sources with Different Contamination Pathways.

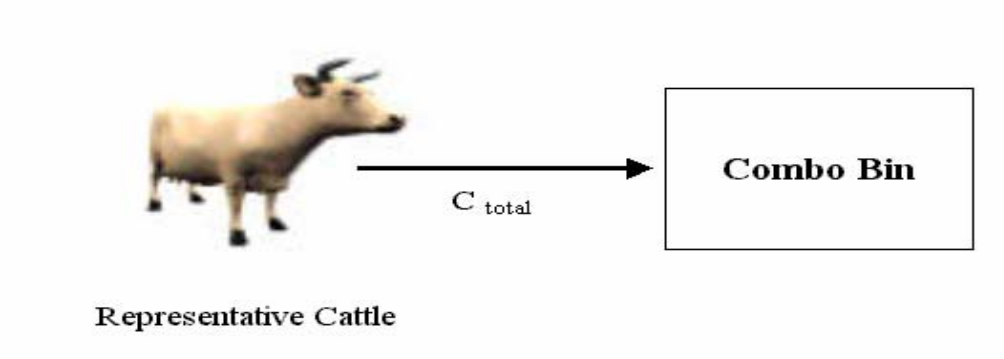



Figure 3-13. The Representative Cattle Introduced in the Modified *E. Coli* Model in the Slaughter Module.

Where,

$OC_i$ = Number of organisms on contaminated carcass i

$TVW_i$ = Trim/Vacuum/Washing efficiency for carcass i (percent)

$EOA_i$ = Evisceration organisms added for carcass i

$WR_i$ = Washing reduction efficiency for carcass i (percent)

$CC_i$ = Number of contaminated $cm^2$ of meat trims for carcass i

$CCM_i$ = Contribution of Each Animal to Combo contamination

### 3.4.3 Preparation Module

Modifications in the preparation module are categorized into two parts: (1) modifications in the serving contamination part; and (2) modifications in the cooking effect part. In the following sections each part is explained separately.

#### 3.4.3.1 Modifications in the Serving Contamination Part

In this section the modifications of the original *E. coli* code in the serving contamination part of the preparation module are explained. In order to clarify the modifications, at first the original approach in calculation of the serving contamination is explained.

In the original *E. coli* model, the amount of ground beef in each food item was calculated using the CSFII recipe files (FSIS, 2001). This provides information about the amount of ground beef consumed during a meal. Consumption data for each ground beef category (e.g., raw meat, hamburger patties, and meatballs) were separated by the eating location (i.e., either at home or away from home). This stratification results in six combinations for ground beef consumption by location: (1) raw ground beef consumed at home; (2) raw ground beef consumed away from

home; (3) ground beef consumed as hamburger at home; (4) ground beef consumed as hamburger away from home; (5) ground beef consumed as meatballs at home; and (6) ground beef consumed as meatballs away from home. Ground beef consumption was further stratified to four age groups: (1) 0 to 5; (2) 6 to 24; (3) 25 to 64; and (4) above 65 years of age. Hence, there are 24 combinations of ground beef consumption type, eating location, and consumer age. For each of these combinations, there is an average serving size based on the CFSII data. In the original code in the serving contamination part, based on the distribution of the grinder load contamination available as an intermediate output of the *E. coli* model in previous parts, the serving contamination distribution is calculated for each combination of ground beef consumption type, eating location and the consumer age, taking in to account the corresponding average serving size. Hence, there are 24 serving contamination distributions for different combinations. Based on the CSFII data, each combination has a total number of servings consumed in the United States. These numbers are implemented to give weight to different combinations. These weights are used to average the serving contamination distributions for different combinations, and finally there is one distribution representing the frequency of contamination in a ground beef serving.

Based on the case scenario for the serving contamination part in Section 3.3.3, the intent is to identify the effect of factors, such as the consumer age, ground beef consumption type, the serving size, and the eating location, on the contamination distribution of a ground beef serving consumed in the United States. In order to achieve this goal, these factors should participate in the simulation as random variables, and not as point estimates or averages as they are behaving in the original *E. coli* model in the serving contamination part. Hence, the original *E. coli* code in the serving contamination part was modified so that these factors participate in the simulation as random variables.

In order to define distributions for these factors, data from CFSII were used. These data are given in Tables 3-5 to 3-7. For these factors discrete distributions were defined, considering the information presented in the tables. For example, in Figure 3-14 the defined discrete distribution for the ground beef consumption type in the United States is depicted. For the consumer age and the eating location, the weights of the defined discrete distributions are given in Tables 3-5 to 3-7. For the serving size, different cumulative distributions are defined considering the data from CFSII survey. The ground beef consumption data from CFSII are

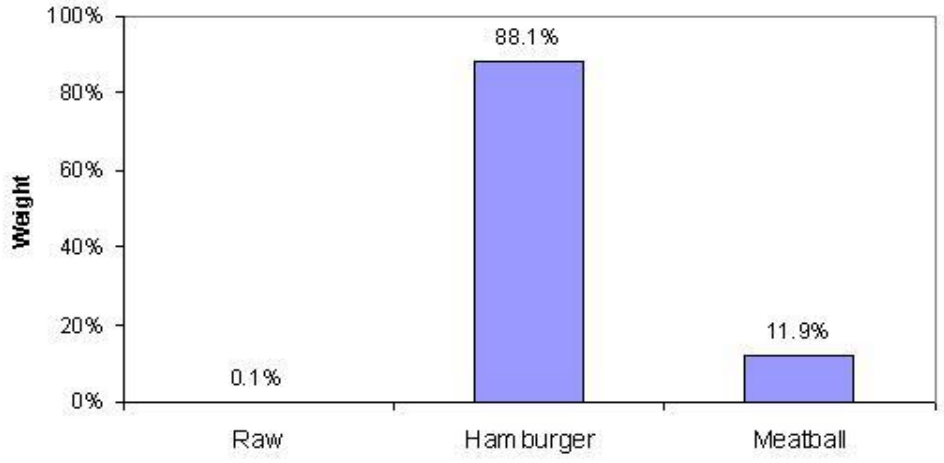



Figure 3-14. Discrete Distribution for the Ground Beef Consumption Type in the US.

available in the original *E. coli* model (i.e., the '*CONSUMPTION*' worksheet). These data are presented in the form of minimum serving size, maximum serving size, and 5 to 95 percentiles of the serving size for each combination of the ground beef consumption type, the consumer age, and the eating location. Hence, 24 cumulative distributions are defined for the serving size.

In Figure 3-15, steps in the modified code for the serving contamination estimation part are depicted. In step 1, a ground beef consumption type is randomly selected from the corresponding discrete distribution. In step 2, the eating location is selected from its distribution. In step 3, the age of the consumer is selected from the corresponding discrete distribution. The selected ground beef consumption type and the eating location are taken into account when choosing a distribution for the consumer age, because for different combinations of the ground beef consumption type and the eating location there are different distributions for consumer age. In step 4, considering selected values for the ground beef consumption type, the eating location, and the consumer age, a cumulative distribution for the serving size is selected from the available 24 distributions, and a serving size is randomly picked from that distribution. In step 5, the grinder load contamination is calculated from the available distribution. In step 6, the mean contamination of the ground beef serving is estimated as the output of interest in this part. Equation 3-7 is used to calculate the mean contamination of the ground beef serving as the output of interest in this part. These steps are repeated for the number of iterations in the simulation.

$$\text{Mean Serving Contamination} = \text{Serving Size} * 10^{\text{Grinder Contamination}} \quad (3\text{-}7)$$

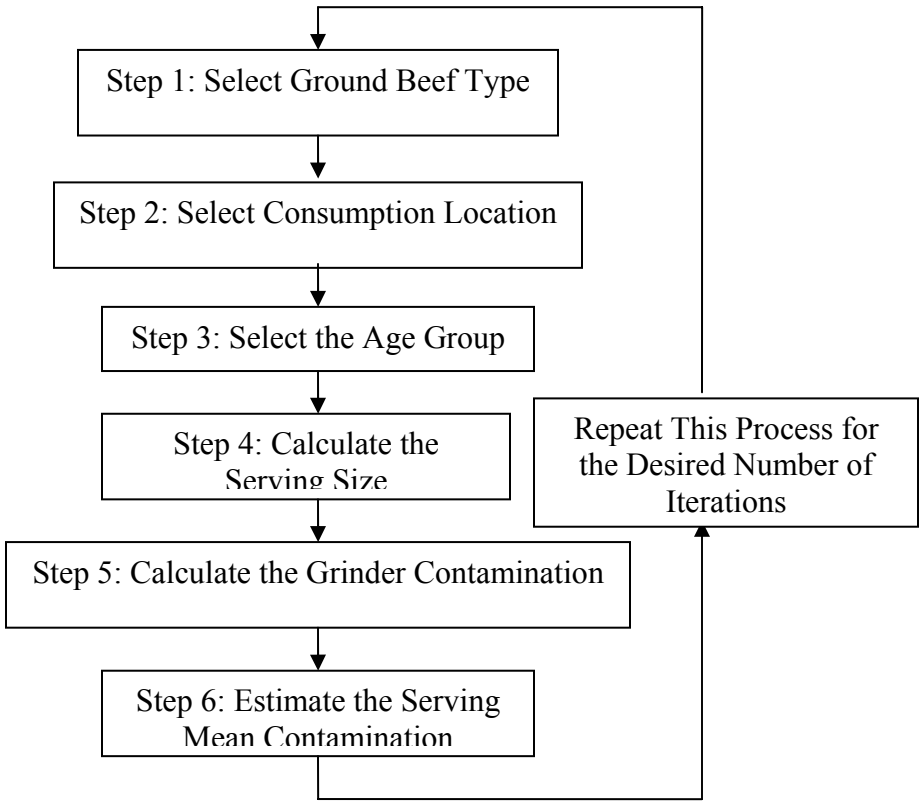



Figure 3-15. Schematic Diagram of Modified Algorithm for the Mean Serving Contamination Calculation.

Table 3-5. Consumption Data for Raw Ground Beef

| Location | Age in Years | Number of Servings | Mean Serving Size (g) | Weight |
|---|---|---|---|---|
| **Home** | 0-5 | ---- | ---- | ---- |
| | 6-24 | ---- | ---- | ---- |
| | 25-64 | 8,861,470 | 113.40 | 100% |
| | 65+ | ---- | ----- | ---- |
| Total | | 8,861,470 | **Weight** | 66.8% |
| **Away** | 0-5 | 522,315 | 56.70 | 11.9% |
| | 6-24 | ---- | ---- | ---- |
| | 25-64 | 3,883,053 | 12.60 | 88.1% |
| | 65+ | ---- | ---- | ---- |
| Total | | 4,405,368 | **Weight** | 33.20% |

Source: *E. coli* food safety risk assessment model, "CONSUMPTION" worksheet

Table 3-6. Consumption Data for Hamburger Patties

| Location | Age in Years | Number of Servings | Mean Serving Size (g) | Weight |
|---|---|---|---|---|
| **Home** | 0-5 | 395,592,840 | 51.86 | 8.0% |
| | 6-24 | 1,478,341,250 | 95.17 | 29.7% |
| | 25-64 | 2,517,532,750 | 102.02 | 50.7% |
| | 65+ | 577,825,295 | 86.52 | 11.6% |
| Total | | 4,969,292,135 | **Weight** | 31.0% |
| **Away** | 0-5 | 717,308,950 | 36.88 | 6.5% |
| | 6-24 | 4,215,244,840 | 78.73 | 38.0% |
| | 25-64 | 5,628,291,058 | 87.64 | 50.8% |
| | 65+ | 523,589,763 | 67.53 | 4.7% |
| Total | | 11,084,434,611 | **Weight** | 69.0% |

Source: *E. coli* food safety risk assessment model, "CONSUMPTION" worksheet

Table 3-7. Consumption Data for Meatballs

| Location | Age in Years | Number of Servings | Mean Serving Size (g) | Weight |
|---|---|---|---|---|
| **Home** | 0-5 | 109,001,410 | 62.36 | 7.6% |
| | 6-24 | 362,621,113 | 123.02 | 25.4% |
| | 25-64 | 686,647,125 | 123.95 | 48.0% |
| | 65+ | 272,269,925 | 100.09 | 19.0% |
| Total | | 1,430,539,573 | **Weight** | 66.2% |
| **Away** | 0-5 | 27,548,375 | 64.01 | 3.8% |
| | 6-24 | 169,672,623 | 75.64 | 23.2% |
| | 25-64 | 398,076,300 | 101.57 | 54.5% |
| | 65+ | 135,376,128 | 67.30 | 18.5% |
| Total | | 730,673,425 | **Weight** | 33.8% |

Source: *E. coli* food safety risk assessment model, "CONSUMPTION" worksheet

### 3.4.4 Modifications in the Cooking Effect Part

In this section the modifications of the original *E. coli* code in the cooking effect part of the preparation module are explained. In order to clarify the modifications, at first the original approach in calculation of the cooking effect on the reduction of the number of *E. coli O157:H7* organisms in ground beef servings is explained.

Cooking effect depends on the cooking temperature, the precooking treatment, and the place of cooking (i.e., home or away from home). Precooking treatment refers to the condition in which a ground beef serving is stored before cooking. There are nine precooking treatments, considered in the *E. coli* food safety risk assessment model. These treatments are summarized in

Table 3-8. Specification of Precooking Treatments in the Cooking Effect Part

| PreCooking Treatment | Specification |
|---|---|
| **A** | $-18^{oC}$ for 8 days |
| **B** | $3^{oC}$ for 9 hours |
| **C** | $-18^{oC}$ for 8 days $21^{oC}$ for 4 hours |
| **D** | $15^{oC}$ for 9 hours |
| **E** | $3^{oC}$ for 9 hours $21^{oC}$ for 4 hours |
| **F** | $-18^{oC}$ for 8 days $30^{oC}$ for 4 hours |
| **G** | $15^{oC}$ for 9 hours $21^{oC}$ for 4 hours |
| **H** | $3^{oC}$ for 9 hours $30^{oC}$ for 4 hours |
| **I** | $15^{oC}$ for 9 hours $30^{oC}$ for 4 hours |

Source: FSIS, 2001

Table 3-8. For each precooking treatment there is a linear regression model, indicating the relation between the cooking temperature and the log reduction in the number of *E. coli O157:H7* organisms due to cooking. A cumulative distribution for the cooking temperature is available. For each precooking treatment a frequency distribution of log reduction in contamination due to cooking is calculated. Hence, at the end of each iteration, there are 9 cooking effect distributions for servings cooked at home and 9 cooking effect distributions for servings cooked away from home. In addition, different weights are allocated to different precooking treatments in each iteration. All of the estimated cooking effect distributions are averaged taking into account these weights. As the final output of the model in this part, there is a single frequency distribution of log reduction in contamination due to cooking.

Based on the case scenario for the cooking effect part in Section 3.3.3, the intent is to identify the effect of precooking treatments and the cooking location on the amount of log reduction in the number of *E. coli O157:H7* organisms due to cooking. In order to address this concern, these factors should participate in the simulation as random variables. Hence, the original *E. coli* code in the cooking effect part was modified in a way that these factors participate in the simulation as random variables.

For the precooking treatment a discrete distribution is defined, using the weight of each treatment as the probability of occurrence of that treatment. Since these weights are changing in different iterations of a random simulation, a range of uncertainty is considered for each precooking treatment probability. Figure 3-16 depicts the discrete distribution defined for the precooking treatment. The depicted probabilities are the mean probabilities of different precooking treatments. In addition, the range of uncertainty regarding the probability of each

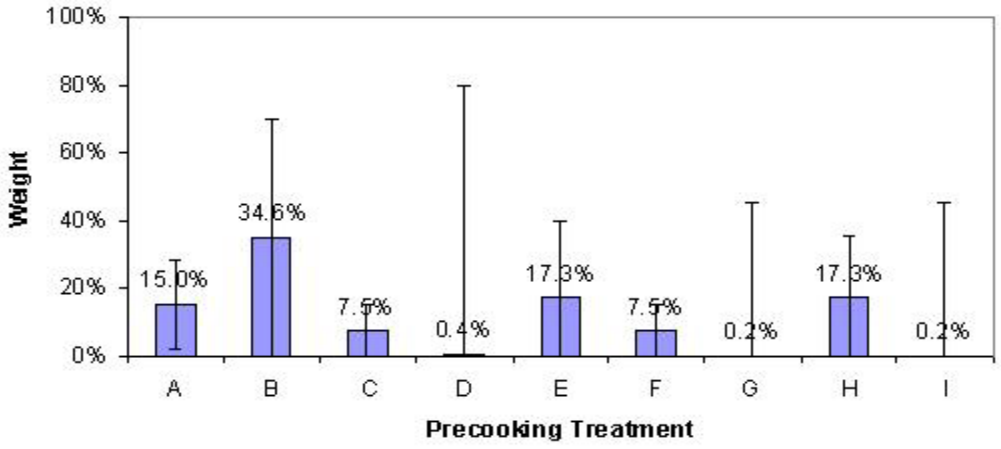



Figure 3-16. Discrete Distribution for the Precooking Treatments Considering the 95% Probability Range for the Mean Values.

precooking treatment is also demonstrated in this figure. For the cooking location, based on information available in the original *E. coli* model regarding the number of ground beef servings consumed at home or away from home, a discrete distribution is defined for this factor.

Figure 3-17 depicts the schematic diagram of the modified algorithm in the cooking effect part**.** In step 1, a precooking treatment is selected from the corresponding distribution. In step 2, a cooking place is selected from the distribution defined for this input. In step 3, a cooking temperature is randomly picked from the cumulative cooking temperature distribution. In step 4, considering the precooking treatment already selected in step 1, the log reduction in the number of *E. coli O157:H7* organisms due to cooking is estimated using the linear regression model available for that treatment. These steps are repeated for the desired number of iterations in the simulation.

## 3.5 Identification of the Input Variables and the Outputs of Interest

In order to perform any sensitivity analysis method a data set containing the paired values of input variables and the outputs of interest are formed. Therefore, it is crucial to identify the input variables and the deserved outputs before performing the simulation, and to modify the code in a way such that those inputs are stored properly. In the *E. coli* model there are input variables in each module affecting the outputs of interest. During the input identification process, it should be clarified whether the selected input represents variability, uncertainty or both. In addition, assumptions made for each input regarding the possible range of variation and its

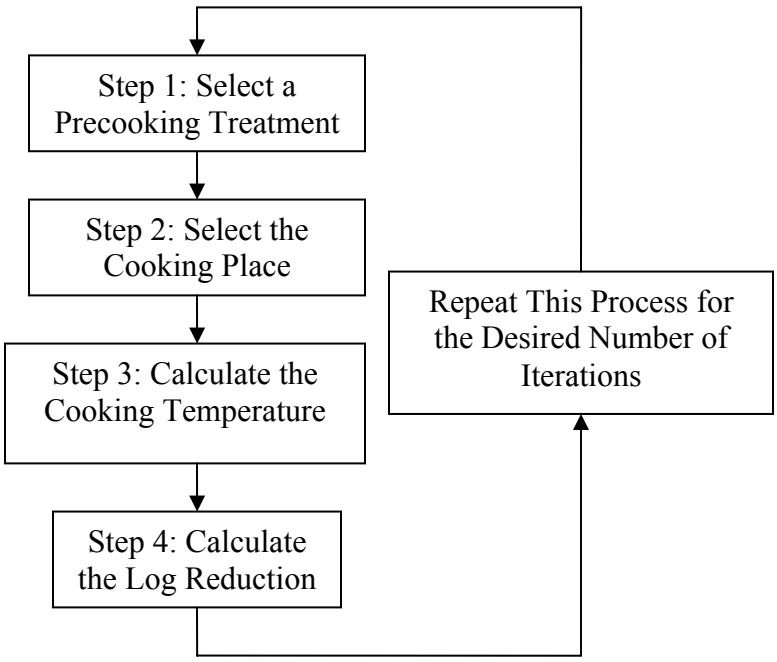



Figure 3-17. Schematic Diagram of Modified Algorithm for the Mean Serving Contamination Calculation.

distribution has to be identified. The input variables and the outputs of interest for each module in the *E. coli* food safety risk assessment model are given in Tables 3-9 to 3-13. In those tables, the "Variable Characteristic" column indicates whether the input represent variability, uncertainty or both.

For inputs incorporating both variability and uncertainty, uncertainty is considered in the parameters of the variability distribution. For example, for the number of combo bins to which each animal contributes, there is a triangular distribution representing the variability of the input. The mean value of this distribution is uncertain. Hence, a uniform distribution is considered for the mean value to represent the uncertainty in this input. In addition, each input can be quantitative or qualitative. A quantitative variable is naturally measured as a number for which meaningful arithmetic operations make sense, while qualitative or categorical variables are any

variable that is not quantitative. Qualitative variables take a value that is one of several possible categories. As naturally measured, categorical variables have no numerical meaning.

Tables 3-9, 3-10, 3-11, 3-12, and 3-13 give the inputs and outputs in the production module, the slaughter module, the growth estimation, the serving contamination, and the cooking effect parts, respectively.

Table 3-9. Input Variables and Outputs of Interest in the Production Module

| Name | Input or Output | Variable Characteristic | Value, Equation, or Distribution | Comment |
|---|---|---|---|---|
| **Feedlot Prevalence** | | | | |
| Study | Input | Uncertain | Discrete[{A,B,C,D};{1,1,1,1}] (1) | Qualitative |
| Apparent Feedlot Prevalence (AFP) | Input | Uncertain | Beta($N_P$+1 , $N_T$-$N_P$+1) (2) | Quantitative |
| Herd Sensitivity | Input | Uncertain | 1-(1-Exponential(AFP))$^{SF}$ (3) | Quantitative |
| Feedlot Prevalence | Output | ---- | Discrete[{$1^{\%}$...$100^{\%}$};{$P_1$...$P_{100}$}] | (4) |
| Within-feedlot Prevalence | | | | |
| Study | Input | Uncertain | Discrete[{A,B,C,D,E},{$P_A$,$P_B$,$P_C$,$P_D$,$P_E$}] (5) | Qualitative |
| Season | Input | Uncertain | Discrete[{Winter, Summer},{1,1}] | Qualitative |
| Apparent Within-feedlot Prevalence (AWFP) | Input | Uncertain | Beta($N_P$+1 , $N_T$-$N_P$+1) (2) | Quantitative |
| Test Sensitivity | Input | Uncertain | Beta($\alpha$ , $\beta$) (6) | Quantitative |
| Average Within-feedlot Prevalence | Output | ---- | Average over high or low prevalence seasons | ---- |
| **Breeding Herd Prevalence** | | | | |
| Study | Input | Uncertain | Discrete[{A,B,C,D,E,F,G};{1,1,1,1,1,1,1}] (1) | Qualitative |
| Apparent Breeding Herd Prevalence (ABHP) | Input | Uncertain | Beta($N_P$+1 , $N_T$-$N_P$+1) (2) | Quantitative |
| Herd Sensitivity | Input | Uncertain | 1-(1-Exponential(ABHP))$^{SF}$ (3) | Quantitative |
| Breeding Herd Prevalence | Output | ---- | Discrete[{$1^{\%}$...$100^{\%}$};{$P_1$...$P_{100}$}] | (4) |
| Within-breeding Herd Prevalence | | | | |
| Study | Input | Uncertain | Discrete[{A,B,C,D,E},{$P_A$,$P_B$,$P_C$,$P_D$,$P_E$}] (5) | Qualitative |
| Season | Input | Uncertain | Discrete[{Winter, Summer},{1,1}] | Qualitative |
| Apparent Within-breeding Herd Prevalence (AWBHP) | Input | Uncertain | Beta($N_P$+1 , $N_T$-$N_P$+1) (2) | Quantitative |
| Test Sensitivity | Input | Uncertain | Beta($\alpha$ , $\beta$) (6) | Quantitative |
| Average Within-breeding Herd Prevalence | Output | ---- | Average over high or low prevalence seasons | ---- |

(Continued on the next page)

Table 3-9. Continued

1. Discrete distributions with equal weights are defined for the study effect in the feedlot and breeding herd prevalence estimation. In the feedlot prevalence, four studies contribute in the estimation of the output, while for the breeding herd prevalence there are seven studies. The data regarding these studies are given in Tables 3-1 and 3-3.
2. $N_P$ = Number of positive cattle in positive herds or feedlots, $N_T$ = Number of cattle tested in positive herds. These values are summarized in Tables 3-1 and 3-3.
3. SF = Samples per feedlot or breeding herd. Different values of SF are given in Tables 3-1 and 3-3 for different studies.
4. The probability of different feedlot or breeding herd prevalence is estimated for different values of prevalence from 1 to 100 percent. In the original *E. coli* model a random value is picked from this distribution as the output in this part, while in the modified model the median value of this discrete distribution is considered as the output of interest.
5. Discrete distributions with unequal weights are defined for the study effect in the within feedlot and breeding herd prevalence estimation. In the within feedlot and breeding herd prevalence parts five studies contribute to the estimation of the output. The data regarding these studies are given in Tables 3-2 and 3-4.
6. Parameters $\alpha$ and $\beta$ differ for different testing methods specific for each study. The testing methods used for each study are specified in Tables 3-2 and 3-4 for feedlots and breeding herds, respectively.

Table 3-10. Input Variables and Outputs of Interest in the Slaughter Module

| Name | Input or Output | Variable Characteristic | Value, Equation, or Distribution | Comment |
|---|---|---|---|---|
| Number of Combo Bins, Each Animal Contributes (Steer/Heifer) | Input | Variable Uncertain | Triangle (2,Uniform (2,5),6) | Quantitative |
| Number of Combo Bins, Each Animal Contributes (Cow/Bull) | Input | Variable Uncertain | Triangle (2, Uniform(2,3),4) | Quantitative |
| Number of Infected Feedlot Cattle in a Lot ($N_{IF}$) | Input | Variable Uncertain | Binomial(1,H).Binomial[40,Exponential(W)] (1) | Quantitative |
| Number of Infected Breeding Herd Cattle in a Lot ($N_{IB}$) | Input | Variable Uncertain | $\sum${Binomial[1,H*Exponential(W)]} (1) | Quantitative |
| Number of Contaminated Feedlot Cattle in a Lot ($N_{CF}$) | Input | Variable Uncertain | Poisson [$N_{IF}$*TR] (2) | Quantitative |
| Number of Contaminated Breeding Herd Cattle in a Lot ($N_{CB}$) | Input | Variable Uncertain | Poisson [$N_{IB}$*TR] (2) | Quantitative |
| Probability of Contamination at Both Steps (Dehiding & Evisceration) ($P_{Both}$) | Input | Variable | Uniform [0,$N_{IF}$ or $N_{IB}$]/($N_{IF}$ or $N_{IB}$) | Quantitative |
| Number of Contaminated Carcasses at Evisceration ($N_{CE}$) | Input | Variable | Binomial [($N_{CF}$ or $N_{CB}$), $P_{Evis}$] (3) | Quantitative |
| Number of Contaminated Carcasses at Evisceration ($N_{CE}$) | Input | Variable | Binomial [($N_{CF}$ or $N_{CB}$), $P_{Evis}$] (3) | Quantitative |
| Number of Contaminated Carcasses at Both Steps (Dehiding & Evisceration) | Input | Variable | Binomial [Min($N_{CF}$ or $N_{CB}$, $N_{CE}$), $P_{Both}$] | Quantitative |
| Chilling Effect | Input | Variable Uncertain | $10^{\text{Normal}\{\text{Uniform}(-0.5,0.5),\,1)}$ | Quantitative |
| Organisms on Contaminated Carcass | Input | Variable Uncertain | I *A (4) | Quantitative |
| Trim/Vacuum/Wash Efficiency | Input | Variable Uncertain | $10^{-\text{Tringular}\{0,\text{Uniform}(0.3,0.7),\text{Uniform}(0.8,1.2)\}}$ | Quantitative |

(Continued on next page)

Table 3-10. Continued

| | | | | |
|---|---|---|---|---|
| Evisceration Organisms Added | Input | Variable Uncertain | I *A [4] | Quantitative |
| Washing Percent Reduction | Input | Variable Uncertain | $10^{-\text{Tringular }\{0,\text{Uniform }(0.5,1.5),\text{Uniform}(1.5,2.5)\}}$ | Quantitative |
| Contaminated $cm^2$ (CC) | Input | Variable Uncertain | Binomial (ζ*φ, A/TSA) [5] | Quantitative |
| Contamination in a Combo Bin | Output | ---- | Poisson $(\eta * CC)$ | (6) |

1. H = Feedlot or breeding herd prevalence, W = Within feedlot or breeding herd prevalence.
2. TR represents the transformation ratio based upon a ratio of two beta distributions. It relates the frequency of contaminated carcasses to the frequency of the infected cattle in a lot.
3. $P_{Evis}$ = Probability of the evisceration occurrence.
4. I = Initial number of organisms on contaminated carcasses introduced during dehiding and is modeled as a cumulative frequency distribution, A = Contaminated surface area.
5. ζ = Weight of contribution, φ = $cm^2$/lb of meat trims, TSA = Total surface area.
6. In the slaughter module the output of interest is calculated using a Poisson distribution with the parameter of the distribution estimated based on the inputs in this module. Hence, in each iteration a value is picked randomly from the distribution as the output of interest. In the modified slaughter module the mean of this Poisson distribution is considered as the output of interest.

Table 3-11. Input Variables and Outputs of Interest in the Preparation Module, Growth Part

| Name | Input or Output | Variable Characteristic | Value, Equation, or Distribution | Comment |
|---|---|---|---|---|
| Storage Temperature Retail | Input | Variable Uncertain | Cumulative | Quantitative |
| Storage Temperature Transportation | Input | Variable Uncertain | Cumulative | Quantitative |
| Storage Temperature Home | Input | Variable Uncertain | Cumulative | Quantitative |
| Storage Time Retail ($ST_1$) | Input | Variable Uncertain | 24 * Exponential[Uniform(0.5,1.5)] | Quantitative |
| Storage Time Transportation ($ST_2$) | Input | Variable | Cumulative | Quantitative |
| Storage Time Home ($ST_3$) | Input | Variable Uncertain | 24 * Exponential[Uniform(0.5,1.5)] | Quantitative |
| Maximum Density | Input | Variable Uncertain | Triangle (5, Uniform(5,10),10) | Quantitative |
| Lag Period Retail | Input | Variable | Exponential{Normal(9.98-2.69Ln($ST_1$),0.27)} | Quantitative |
| Lag Period Transportation | Input | Variable | Exponential{Normal(9.98-2.69Ln($ST_2$),0.27)} | Quantitative |
| Lag Period Home | Input | Variable | Exponential{Normal(9.98-2.69Ln($ST_3$),0.27)} | Quantitative |
| Generation Time Retail | Input | Variable Uncertain | Exponential{Normal(9.98-2.69Ln($ST_1$),0.16)} | Quantitative |
| Generation Time Transportation | Input | Variable Uncertain | Exponential{Normal(9.98-2.69Ln($ST_2$),0.16)} | Quantitative |
| Generation Time Home | Input | Variable Uncertain | Exponential{Normal(9.98-2.69Ln($ST_3$),0.16)} | Quantitative |
| Growth | Output | ---- | ----- | ---- |

Table 3-12. Input Variables and Outputs of Interest in the Preparation Module, Serving Contamination Part

| Name | Input or Output | Variable Characteristic | Value, Equation, or Distribution | Comment |
|---|---|---|---|---|
| Ground Beef Consumption Type | Input | Variable | Discrete[{Raw,Hamburger,Meatball},{$0.1^{\%}$,$88^{\%}$,$12^{\%}$}] | Qualitative |
| Eating Location | Input | Variable | Discrete[{Home, Away},{$P_H$,$P_W$}] | Qualitative |
| Age of Consumer | Input | Variable | Discrete[{A,B,C,D,E},{$P_A$,$P_B$,$P_C$,$P_D$,$P_E$}] | Qualitative |
| Grinder Contamination (GC) | Input | Variable | Cumulative | Quantitative |
| Serving Size (SS) | Input | Variable | Cumulative | Quantitative |
| Serving Contamination | Output | ---- | Poisson (SS * $10^{GC}$) | (1) |

1. In the serving contamination part, the output of interest is the contamination in a ground beef serving that is estimated by randomly picking a value from a Poisson distribution with the distribution parameter estimated using inputs in this part. In the modified model, the mean value of this distribution is considered as the output of interest.

Table 3-13. Explanatory Variables and Outputs of Interest in the Preparation Module, Cooking Effect Part

| Name | Input or Output | Variable Characteristic | Value, Equation, or Distribution | Comment |
|---|---|---|---|---|
| Precooking Treatment | Input | Variable | Discrete[{A,…,I},{$P_A$,…,$P_I$}] | Qualitative |
| Cooking Place | Input | Variable | Discrete[{Home, Away},{0.35, 0.65}] | Qualitative |
| Cooking Temperature | Input | Variable | Beta($\alpha$ , $\beta$) | Quantitative |
| Log Reduction | Output | ---- | Intercept + Slope * T | (1) |

1. The values of the intercept and slope of the linear log reduction model for each precooking treatment are specified separately in the model.

# 4 NOMINAL RANGE SENSITIVITY ANALYSIS APPLIED TO THE *E. COLI* MODEL

The objective of this chapter is to evaluate the application of NRSA to the *E. coli* model. This chapter includes three sections. Section 4.1 explains limitations regarding the application of NRSA to the ***E. coli*** model. Section 4.2 presents an example to illustrate NRSA application to the growth estimation part of the preparation module. Section 4.3 presents a summary and conclusion regarding the application of the mathematical sensitivity analysis methods to the *E. coli* model.

## 4.1 Overview of Limitations Precluding the Application of NRSA to the *E. coli* Model

This section explains the reasons that make the application of NRSA to the *E. coli* model impractical. These reasons are classified into two categories: (1) model limitation; and (2) method limitation.

Section 3.2.4.1 explained that one of the limitations of the *E. coli* model for application of sensitivity analysis methods is that the model is structured in Microsoft Excel using inter-cellular functioning. The use of equations inside worksheet cells, as opposed to the use of a stand-alone programming language, permits execution of the code in the spreadsheet environment; however, spreadsheet-based models are difficult to modify compared to programming language-based models. In order to apply NRSA, all inputs should be held at nominal point values, and only one input is varied to its minimum and maximum values to evaluate the effect of this variation on the output. Because all the model equations are stored in worksheet cells, changing the values of cells is difficult to implement.

In Section 2.1.1 NRSA is explained. NRSA is a simple sensitivity analysis method that gives insight regarding the relative effect of inputs on the output change. Since NRSA is not prepared basically to address special relationship such as interactions between inputs and nonlinearity in the model response, when there are such characteristics in the model NRSA may not present informative results. In those cases that there are interactions between inputs, NRSA may give results if changing each input value is synchronized with change in the input that has interaction with the first input. This process may be tedious and time consuming, because at the beginning of the analysis it is not clear as to which inputs have interaction. Hence, it is possible that too many combinations would have to be examined in order to find those inputs that have

interaction. The amount of manual work that has to be done increases substantially when the number of inputs to the model increases.

Statistical and graphical methods for the sensitivity analysis applied in Chapters 5 to 9 indicate that there are statistically significant interactions between inputs to different parts and modules of the *E. coli* model. Thus, considering the limitations explained above, NRSA was not selected for application to the *E. coli* model.

In next section an example is presented for application of NRSA to the growth estimation part. Through this example, it is illustrated that how an interaction effect between the storage time and temperature affects the results from NRSA.

## 4.2 Example of Application of NRSA to the *E. coli* Model

In this section an example is presented regarding the application of NRSA in the growth estimation part of the production module. In addition to the simple NRSA, the possible interaction between inputs was addressed by conditioning the change in the input values to changes in the values of other inputs that are suspected to have an interaction with the first input.

Table 3-11 summarizes the inputs and their distributions in the growth estimation part. Nominal values (i.e. minimum, mean, and maximum) for each input are extracted considering the input distribution. These values are given in Table 4.1.

Equation 2-1 is used as an index for the sensitivity measurement. In Table 4-2 the results of simple NRSA method are summarized. Since the model response when all the variables are at mean values is zero (i.e. no growth), the sensitivity indexes based on that equation could not be estimated. Therefore only the numerator of the equation was only used for measuring the sensitivity.

Results in Table 4-2 implies that for almost all the inputs changing the input value between the maximum and minimum values, while conditioning other inputs at their mean values does not result in any change in the model response. Thus, there may be interactions between inputs. In each stage, growth is estimated if the available time at that stage is longer than the lag period of the stage, otherwise the growth will be estimated as zero (FSIS, 2001). Mean values of the storage times and the lag periods at different stages in Table 4-1 indicate that for the case with all the inputs conditioned at their mean values the growth is zero. Moreover, the growth in the number of *E. coli* organisms in ground beef servings could be estimated only if the storage time and the lag period are changed simultaneously.

Table 4-1. Nominal Values for Inputs to the Growth Estimation Part

| Variable | Minimum | Mean | Maximum | Unit |
|---|---|---|---|---|
| Storage Time at Retail | 0 | 24 | 340 | Hour |
| Storage Temperature at Retail | 46 | 47.6 | 73 | $^{o}$F |
| Storage Time at Transportation | 0 | 1 | 6.5 | Hour |
| Storage Temperature at Transportation | 46 | 48.8 | 73 | $^{o}$F |
| Storage Time at Home | 0 | 24 | 340 | Hour |
| Storage Temperature at Home | 46 | 48.3 | 73 | $^{o}$F |
| Maximum Density | 5 | 7.5 | 10 | log |
| Lag Period at Retail | 2.7 | 73.6 | 250.2 | Hour |
| Lag Period at Transportation | 2.5 | 64.7 | 260.9 | Hour |
| Lag Period at Home | 2.2 | 71.8 | 247.5 | Hour |
| Generation Time at Retail | 0.6 | 9.9 | 22.8 | Hour |
| Generation Time at Transportation | 0.6 | 8.7 | 24.6 | Hour |
| Generation Time at Home | 0.6 | 9.7 | 24 | Hour |

Table 4-2. The Results of the NRSA in the Growth Estimation Part

| Variable | Growth in Log at Input Values: | | | Sensitivity (Rank) |
|---|---|---|---|---|
| | Minimum | Mean | Maximum | |
| Storage Time at Retail | 0 | 0 | 0.7 | 0.7(1) |
| Storage Temperature at Retail | 0 | 0 | 0 | 0 |
| Storage Time at Transportation | 0 | 0 | 0 | 0 |
| Storage Temperature at Transportation | 0 | 0 | 0 | 0 |
| Storage Time at Home | 0 | 0 | 0.7 | 0.7(1) |
| Storage Temperature at Home | 0 | 0 | 0 | 0 |
| Maximum Density | 0 | 0 | 0 | 0 |
| Lag Period at Retail | 0.09 | 0 | 0 | -0.09(2) |
| Lag Period at Transportation | 0 | 0 | 0 | 0 |
| Lag Period at Home | 0.07 | 0 | 0 | -0.07(3) |
| Generation Time at Retail | 0 | 0 | 0 | 0 |
| Generation Time at Transportation | 0 | 0 | 0 | 0 |
| Generation Time at Home | 0 | 0 | 0 | 0 |

In order to incorporate the effect of interactions in NRSA, conditional NRSA was applied to this part. Table 4-3 gives the results of conditional NRSA at stage one (i.e. retail). Although just three inputs are considered in the example, 27 manual calculations were performed in order to estimate the sensitivity indexes for inputs.

Table 4-3. The Results of the Conditional NRSA for the Stage One (Retail) in the Growth Estimation Part

| | Storage Time at Retail | Lag Period at Retail | Growth in Log | Sensitivity (Rank) |
|---|---|---|---|---|
| **Storage Temperature at Retail (Low Value)** | Low Value | Low Value | 0 | 0 |
| | | Mean Value | 0 | |
| | | High Value | 0 | |
| | Mean Value | Low Value | 0.74 | 0.74 (4) |
| | | Mean Value | 0 | |
| | | High Value | 0 | |
| | High Value | Low Value | 7.19 | 4.43 (1) |
| | | Mean Value | 7.19 | |
| | | High Value | 2.76 | |
| **Storage Temperature at Retail (Mean Value)** | Low Value | Low Value | 0 | 0 |
| | | Mean Value | 0 | |
| | | High Value | 0 | |
| | Mean Value | Low Value | 0.74 | 0.74 (4) |
| | | Mean Value | 0 | |
| | | High Value | 0 | |
| | High Value | Low Value | 7.15 | 4.4 (2) |
| | | Mean Value | 7.15 | |
| | | High Value | 2.75 | |
| **Storage Temperature at Retail (Maximum Value)** | Low Value | Low Value | 0 | 0 |
| | | Mean Value | 0 | |
| | | High Value | 0 | |
| | Mean Value | Low Value | 0.74 | 0.74 (4) |
| | | Mean Value | 0 | |
| | | High Value | 0 | |
| | High Value | Low Value | 6.59 | 3.83 (3) |
| | | Mean Value | **6.59** | |
| | | High Value | **2.76** | |

Results in Table 4.3 indicate that for low storage temperature at retail, the model response is greater than zero only if there is a low value of lag at a mean storage time or for any lag for a high value of storage time. Thus, it is clear that interactions are important and that the model is non-linear with possible thresholds.

If all three stages (i.e. retail, transportation and home) were considered simultaneously, the number of calculations with considering the interactions between inputs would be boosted dramatically making the analysis onerous and time consuming.

## 4.3 Summary and Conclusions Regarding Application of Mathematical Methods for Sensitivity Analysis to the *E. coli* Model

In Section 4.1 it was explained that because of limitations of the model regarding the modeling environment and those of the NRSA, application of NRSA to the *E. coli* model is not practical or informative. Section 4.2 presented an example for application of NRSA to the growth estimation part. In that example, an attempt was made to address the interaction effect using NRSA. As demonstrated in the example, when the number of inputs to the model increases the application of the conditional NRSA in order to address the interaction effects between inputs becomes impractical. Therefore, it was decided that this method would not be applied to different modules and parts of the model. Application of NRSA is presented with *Listeria monocytogenes* model in Chapter 13, in which there is limited number of inputs to the model.

Moreover, in Section 2.1.2 DSA was explained as another mathematical approach for the sensitivity analysis is explained. This method has almost the same characteristics as the NRSA. Therefore, based on the discussion in this chapter for refraining the application of NRSA, DSA was not applied to the *E. coli* model.

# 5 ANALYSIS OF VARIANCE FOR THE *E. COLI* O157:H7 MODEL

The objective of this chapter is to present the results of applying ANOVA, as a method for the sensitivity analysis, to different modules and parts of the *E. coli* food safety risk assessment model. The modules and parts are explained in Sections 3.2.1 to 3.2.3. ANOVA is explained in Section 2.2.2, including terms specific to ANOVA such as factor, treatment, level, balanced or unbalanced experiment, contrasts, and F values.

In Section 5.1, the specification of levels for different factors affecting the output of interest in each module is explained. In Sections 5.2 to 5.4 the results of ANOVA are presented for the three major modules of the *E. coli* model: (1) production; (2) slaughter; and (3) preparation. For each module, the approach employed regarding the consideration of variability and uncertainty is explained. Details for scenarios for each module are presented in Sections 3.3.1 to 3.3.3.

As an illustration of a technique for performing a diagnostic check on the results of ANOVA, three case studies are provided in Sections 5.4.1.1 to 5.4.1.3 in which the coefficient of determination, $R^2$, is calculated based upon the results of ANOVA. Although the F values calculated for each effect indicate the statistical significance of the respective effect, the coefficient of determination provides insight regarding whether the selected effects adequately capture variability. Moreover, a high value of $R^2$ implies that results are not compromised by inappropriate definition of the levels for each factor.

A case study is provided in Section 5.4.2 to evaluate the uncertainty in F values as an index of sensitivity in ANOVA. The purpose of this case study is to evaluate how large differences must be between F values in order to discriminate the importance of two or more inputs.

In Section 5-5, the ANOVA method is evaluated and the limitations, advantages, disadvantages and key criteria for application of this method to sensitivity analyses are summarized.

## 5.1 Identification of Levels of Factors in Different Modules of the *E. coli* Model

In this section the levels assigned to factors for different modules of the *E. coli* model are identified. Each factor included in ANOVA must be divided into discrete levels. For continuous factors, levels were defined by dividing the domain of values into ranges based upon the cumulative distribution function (CDF) of the factor. In particular, levels are defined based upon

Table 5-1. Levels for Factors Used in ANOVA in the Feedlot Prevalence Part

| Factor | Number of levels | Levels and Corresponding Percentiles (1) | Units |
|---|---|---|---|
| Study | 4 | Dargatz Hancock 1997, Hancock 1998, Smith 1999, Elder 2000 (2) | --- |
| Apparent Prevalence | 3 | {0-4, 4-35, >35}<br>{40th, 80th} Percentiles | Percent |
| Herd Sensitivity | 3 | {0-70, 70-95, >95}<br>{20th, 40th} Percentiles | Percent |

(1) For continuous variables, the ranges that define each factor level and the percentile of the CDF corresponding to the break point between factor levels are given. For discrete variables, each factor level is identified.
(2) Data regarding each study are summarized in Table 3-1.

Table 5-2. Levels for Factors Used in ANOVA in the Within Feedlot Prevalence Part

| Factor | Number of levels | Levels and Corresponding Percentiles (1) | Units |
|---|---|---|---|
| Study | 5 | Dargatz Hancock 1997, Hancock 1999, Hancock 1998, Smith 1999, Elder 2000 (2) | --- |
| Season | 2 | Summer, Winter | --- |
| Apparent Within Feedlot Prevalence | 2 | {0-6, >6}<br>60th Percentile | Percent |
| Test Sensitivity | 4 | {0-50, 50-65, 65-92, >92}<br>{20th, 60th, 80th} Percentiles | Percent |

(1) For continuous variables, the ranges that define each factor level and the percentile of the CDF corresponding to the break point between factor levels are given. For discrete variables, each factor level is identified.
(2) Data regarding each study are summarized in Table 3-1.

the lower tail, middle region, and upper tail of the distribution of each factor. The CDF for each factor is derived based on the generated values from the corresponding distribution during a random simulation of the *E. coli* model. In cases with contributions of both variability and uncertainty in the simulation, levels are identified based on the cumulative distributions developed from co-mingled variability and uncertainty generated values. In Sections 5.1.1 to 5.1.3 these levels are identified for factors in the three modules of the *E. coli* model.

### 5.1.1 Production Module

The production module includes four parts: (1) the feedlot prevalence; (2) the breeding herd prevalence; (3) the within feedlots prevalence; and (4) the within breeding herds prevalence. These parts are explained in Section 3.2.1. In Section 3.3.2 the case scenario for the production module is explained. Factors in the production module are summarized in Table 3-9. Based on the case scenario in the production module, the analysis includes one dimensional uncertainty simulation with 65,000 iterations. CDFs for factors in the production module were derived based

Table 5-3. Levels for Factors Used in ANOVA in the Breeding Herd Prevalence Part

| Factor | Number of levels | Levels and Corresponding Percentiles (1) | Units |
|---|---|---|---|
| Study | 7 | Garber 1998, Sargeant 2000, Hancock/FDA 2001, Hancock 1997a, Hancock 1998, Lagreid 1998, Hancock 1997b (2) | --- |
| Apparent Prevalence | 2 | {0-3, >3}<br>60th Percentile | Percent |
| Herd Sensitivity | 3 | {0-75, 75-94, >94}<br>{20th, 40th} Percentiles | Percent |

(1) For continuous variables, the ranges that define each factor level and the percentile of the CDF corresponding to the break point between factor levels are given. For discrete variables, each factor level is identified.
(2) Data regarding each study are summarized in Table 3-1.

Table 5-4. Levels for Factors Used in ANOVA in the Within Breeding Herd Prevalence Part

| Factor | Number of levels | Levels and Corresponding Percentiles (1) | Units |
|---|---|---|---|
| Study | 6 | Garber 1998, Besser 1997, Rice 1997, Hancock 1994, Sargeant 2000, Hancock/ FDA 2001 (2) | --- |
| Season | 2 | Summer, Winter | --- |
| Apparent Within Breeding Herd Prevalence | 2 | {0-2, >2}<br>50th Percentile | Percent |
| Test Sensitivity | 3 | {0-50, 50-70, >70}<br>{20th, 80th} Percentiles | Percent |

(1) For continuous variables, the ranges that define each factor level and the percentile of the CDF corresponding to the break point between factor levels are given. For discrete variables, each factor level is identified.
(2) Data regarding each study are summarized in Table 3-1.

on the generated values from the uncertainty simulation. These CDFs are depicted in Figure 5-1 for apparent prevalence, herd sensitivity, test sensitivity, apparent within feedlot prevalence, and apparent within breeding herd prevalence.

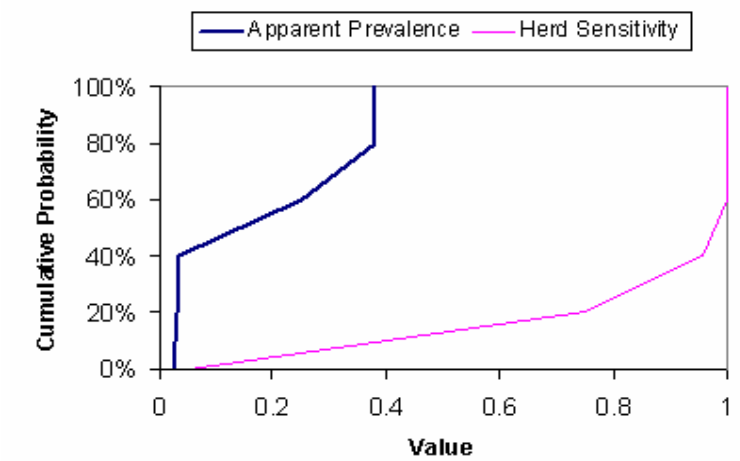
The shape of the CDF for each factor is used to define the levels. For example, for the apparent prevalence in the feedlot prevalence part, shown in Figure 5-1(a), values at the 40th and 80th percentiles are used to define the levels. At these percentiles, the corresponding CDF graph for the apparent prevalence shows changes in the trend of the graph. Hence, these percentiles are selected in order to define the levels for the apparent prevalence. For qualitative factors such as the study and the season, original values generated from the distribution are considered as

different levels. For example, each study is treated as a different level of the study factor. In Tables 5-1 to 5-4 levels for the factors in the production module are summarized.

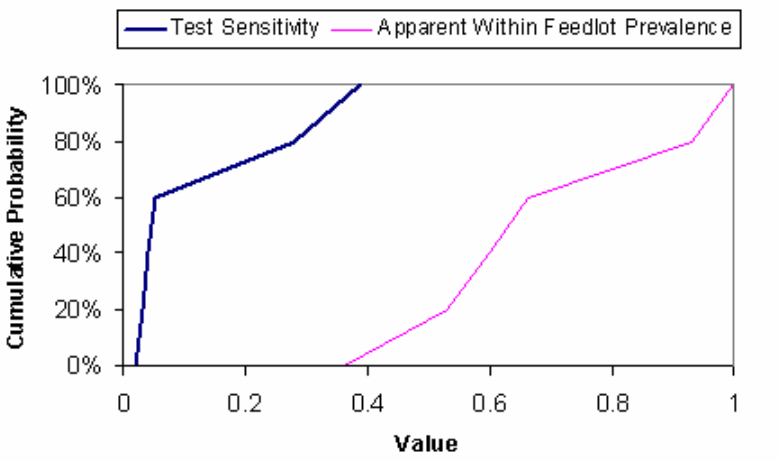
### 5.1.2 Slaughter Module

The slaughter module is explained in Section 3.2.2, including the case scenario. The slaughter module has a two-dimensional variability and uncertainty simulation. Factors for the slaughter module are summarized in Table 3-10. CDFs for some factors in the slaughter module are derived based on the generated values in the co-mingled variability and uncertainty simulation. These CDFs are depicted in Figure 5-2. In the co-mingled variability and uncertainty simulation there are 650 iterations for variability and 100 iterations for uncertainty for a total of 65,000 iterations. Factor levels were defined based on the shape of the corresponding CDF for each factor. For example, for the total number of infected animals, shown in Figure 5-2(a), values at the $25^{th}$, $40^{th}$, and $80^{th}$ percentiles are used to define four levels.

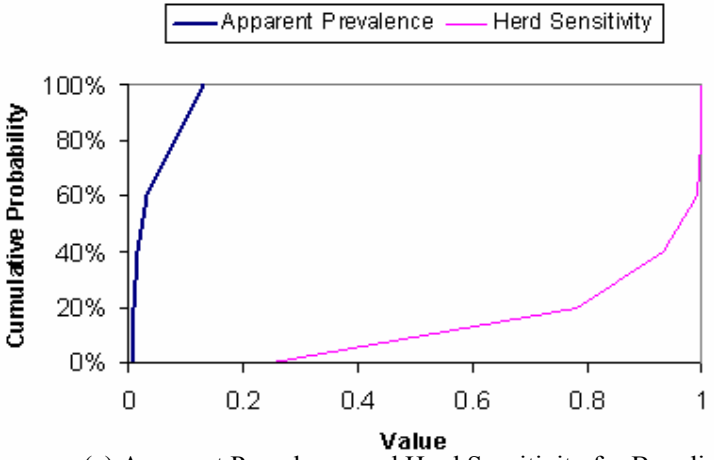
For factors not depicted in Figure 5-2, each unique generated factor value during the simulation was considered as the factor level. For example, the number of combo bins to which each carcass contributes is a factor in the slaughter module. Each possible integer value of the number of combo bins to which each carcass contributes is defined as a factor level. Hence, there are five levels for this factor since a carcass is assumed to contribute to as few as two but not more than six combo bins. For the chilling effect five levels are defined based upon one logarithmic range increments. In Table 5-5, levels for different factors in the slaughter module are summarized.



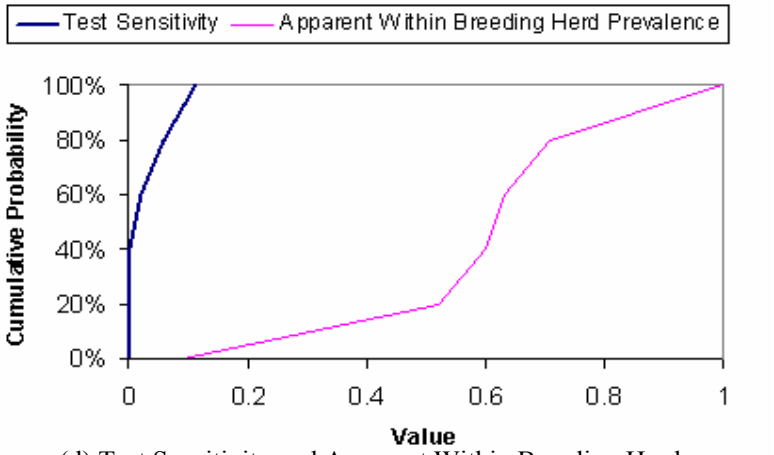

(a) Apparent Prevalence and Herd Sensitivity for Feedlot Prevalence




(b) Test Sensitivity and Apparent Within Feedlot Prevalence for Within Feedlot Prevalence




(c) Apparent Prevalence and Herd Sensitivity for Breeding Herd Prevalence



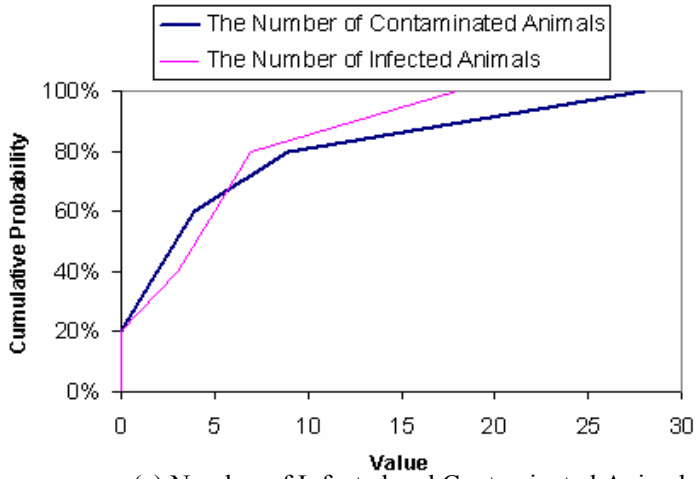

(d) Test Sensitivity and Apparent Within Breeding Herd Prevalence for Within Breeding Herd Prevalence

Figure 5-1. Cumulative Probability Functions for Apparent Prevalence, Herd Sensitivity, Test Sensitivity, and Apparent Within Feedlot and Breeding Herd Prevalence.

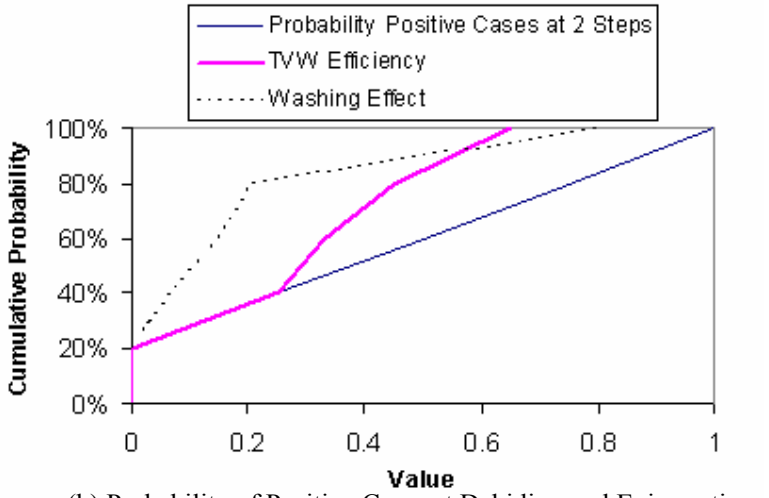
Table 5-5. Levels for Factors Used in ANOVA in the Slaughter Module

| Factor | Number of levels | Levels and Corresponding Percentiles [1] | Units |
|---|---|---|---|
| Total Number of Combo Bins for Each Carcass | 5 | 2, 3, 4, 5, 6 [2] | --- |
| Total Number of Infected Animals | 4 | {0-2, 2-4, 4-6, >6}<br>{25th, 40th, 80th} Percentiles | --- |
| Total Number of Contaminated Animals | 4 | {0-1, 1-4, 4-8, >8}<br>{25th, 60th, 80th} Percentiles | --- |
| Probability of Positive Cases at Dehiding & Evisceration Steps | 4 | 0-25, 25-50,50-75, >75 [3] | Percent |
| Number of Positive Cases at Dehiding & Evisceration Steps | 4 | 0, 1, 2, 3 [2] | --- |
| Number of Positive Cases at Evisceration | 4 | 0, 1, 2, 3 [2] | --- |
| Chilling Effect | 5 | <-1, -1-0, 0-1, 1-2, >2 | Log |
| Number of Organisms | 4 | {0-5, 5-20, 20-70, >70}<br>{50th, 60th, 80th} Percentiles | --- |
| Trim/Vacuum/Washing Efficiency | 3 | {0-25, 25-40, >40}<br>{40th, 80th} Percentiles | Percent |
| Evisceration Organisms Added | 4 | {0-5, 5-20, 20-70, >70} [4]<br>{50th, 60th, 80th} Percentiles | --- |
| Washing Effect | 3 | {0, 0-20, >20}<br>{20th, 80th} Percentiles | Percent |
| Contaminated $cm^2$ | 4 | {0-60, 60-200, 200-600, >60}<br>{40th, 60th, 80th} Percentiles | --- |

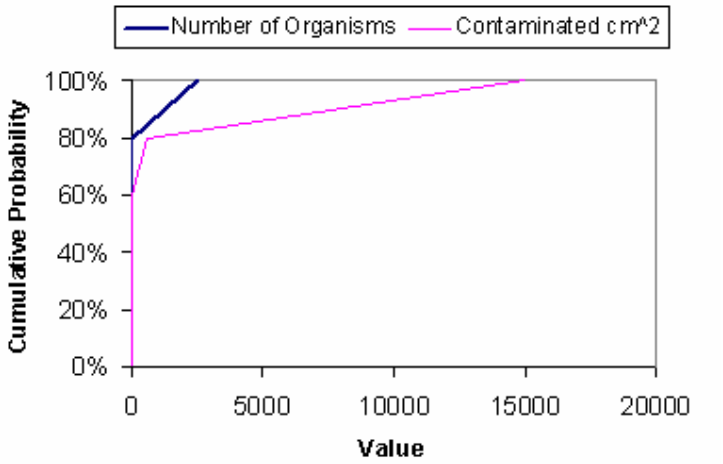
(1) For continuous variables, the ranges that define each factor level and the percentile of the CDF corresponding to the break point between factor levels are given. For discrete variables, each factor level is identified.
(2) Levels are identified based on the generated values for the factor.
(3) For this factor four levels with equal intervals are defined.
(4) Similar levels as the number of organisms are defined for this factor.




(a) Number of Infected and Contaminated Animals




(b) Probability of Positive Cases at Dehiding and Evisceration, TVW Efficiency, and Washing Effect




(c) Number of Organisms and Number of Contaminated $cm^2$ of Meat Trims

Figure 5-2. Cumulative Probability Functions for Factors in the Slaughter Module.

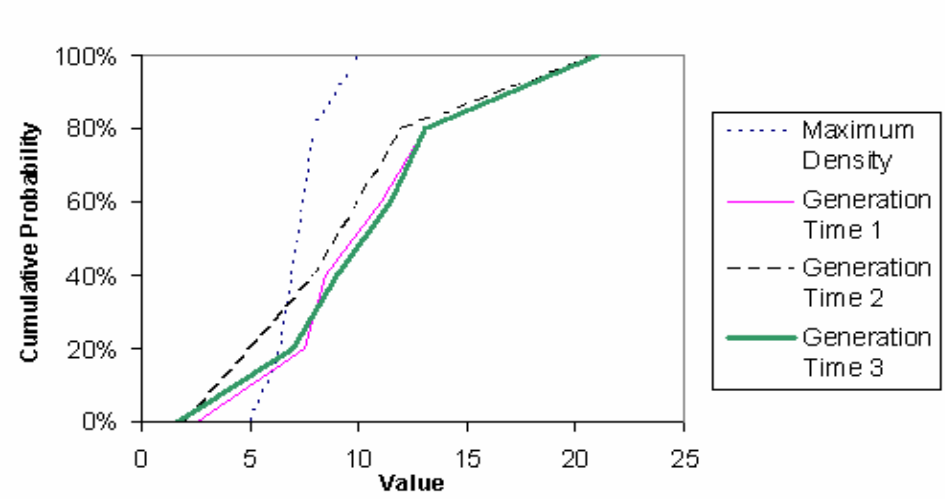



Figure 5-3. Cumulative Probability Functions for the Maximum Density and the Generation Time at Three Stages in the Growth Estimation Part.

### 5.1.3 Preparation Module

The preparation module includes three parts: (1) growth estimation; (2) cooking effect; and (3) serving contamination. These parts are explained in Section 3.2.1. In Section 3.3.2 the case scenarios for the preparation module are explained. Factors for the preparation module are summarized in Table 3-11. In the growth estimation part there is a two-dimensional simulation of variability and uncertainty, while in the cooking effect and the serving contamination parts there is a one-dimensional variability simulation.

In order to identify levels for the factors, CDFs were developed using the simulated values of each factor in a co-mingled variability and uncertainty simulation in the growth estimation part, while for the cooking effect and the serving contamination parts values generated in the variability simulation were used. These CDFs were used only to define levels for quantitative factors. Factor levels were defined based on the shape of the corresponding CDF for each factor. For example, for the maximum density in the growth estimation part, shown in Figure 5-3, values at the $20^{th}$, $50^{th}$, and $80^{th}$ percentiles were used to define four levels. For the qualitative factors, discrete values were used as levels. For example, the cooking place is a qualitative factor with two levels: (1) home; and (2) away from home. Table 5-6 summarizes the levels for factors in the cooking effect and the serving contamination part of the preparation module. Table 5-7 summarizes the levels for factors in the growth estimation part of the preparation module. Figures 5-3 and 5-4 present the CDFs for maximum density, generation times, and lag periods in stages 1 to 3 in the growth estimation part.

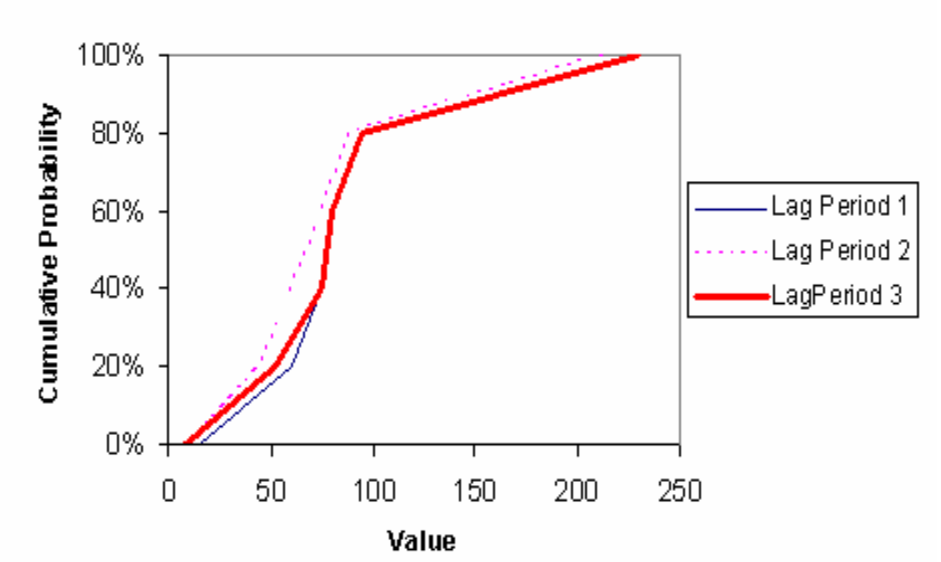



Figure 5-4. Cumulative Probability Functions for the Lag Period at Three Stages in the Growth Estimation Part.

Table 5-6. Levels for Factors Used in ANOVA in the Preparation Module, the Cooking Effect and Serving Contamination Part

| Factor | Number of levels | Levels and Corresponding Percentiles (1) | Units |
|---|---|---|---|
| **Cooking Effect Part** | | | |
| Precooking Treatment | 9 | A, B, C, D, E, F, G, H, I (2) | --- |
| Cooking Place | 2 | Home, Away (1) | --- |
| Cooking Temperature | 5 | {39-58, 58-66, 66-73, 73-79, >79} {20$^{th}$, 40$^{th}$, 60$^{th}$, 80$^{th}$} Percentiles (3) | $^{o}$C |
| **Serving Contamination Part** | | | |
| Ground Beef Consumption Type | 3 | Raw, Hamburger, Meatball | --- |
| Eating Location | 2 | Home, Away | --- |
| Consumer Age | 4 | <5, 5-24, 25-64, >65 | year |
| Serving Size | 7 | 0-30, 30-60, 60-90, 90-120, 120-150, 150-180, >180 (4) | g |
| Grinder Contamination | 7 | <-6, (-6)-(-5), (-5)-(-4), (-4)-(-3), (-3)-(-2), (-2)-(-1), >-1 (4) | log |

(1) For continuous variable, the range that define each factor level and the percentile of the CDF corresponding to the break point between factor levels are given. For discrete variables, each factor level is identified
(2) Levels for the precooking treatment were defined in Table 3-8.
(3) For the cooking temperature equal percentiles are considered as levels.
(4) For this factor equal intervals are considered as levels.

Table 5-7. Levels for Factors Used in ANOVA in the Preparation Module, the Growth Estimation Part

| Factor | Number of levels | Levels and Corresponding Percentiles (1) | Units |
|---|---|---|---|
| Storage Temperature, Stage 1 | 5 | 7.5-11, 11-14.5, 14.5-18, 18-21.5, >21.5 (1) | $^{o}$C |
| Storage Temperature, Stage 2 | 3 | 7.5-13.5, 13.5-19.5, >19.5 (1) | $^{o}$C |
| Storage Temperature, Stage 3 | 5 | 7.5-11, 11-14.5, 14.5-18, 18-21.5, >21.5 (1) | $^{o}$C |
| Storage Time, Stage 1 | 12 | 0-24, 24-48, …, 264-288, >288(2) | hr |
| Storage Time, Stage 2 | 2 | 0-3.5, >3.5 | hr |
| Storage Time, Stage 3 | 12 | 0-24, 24-48, …, 264-288, >288(2) | hr |
| Maximum Density | 4 | {<6.5,6.5-7.5, 7.5-8.5, >8.5} {20$^{th}$, 50$^{th}$, 80$^{th}$} Percentiles | log |
| Lag Period, Stage 1 | 4 | {<50, 50-65, 65-95, >95} {20$^{th}$, 50$^{th}$, 80$^{th}$} Percentiles | hr |
| Lag Period, Stage 2 | 4 | {<35, 35-55, 55-90, >90} {20$^{th}$, 50$^{th}$, 80$^{th}$} Percentiles | hr |
| Lag Period, Stage 3 | 4 | {<45, 45-65, 65-95, >95} {20$^{th}$, 50$^{th}$, 80$^{th}$} Percentiles | hr |
| Generation Time, Stage 1 | 4 | {<7, 7-9.5, 9.5-12.5, >12.5} {20$^{th}$, 50$^{th}$, 80$^{th}$} Percentiles | hr |
| Generation Time, Stage 2 | 4 | {<4.5, 4.5-8, 8-12, >12} {20$^{th}$, 50$^{th}$, 80$^{th}$} Percentiles | hr |
| Generation Time, Stage 3 | 4 | {<6.5, 6.5-9.5, 9.5-13, >13} {20$^{th}$, 50$^{th}$, 80$^{th}$} Percentiles | hr |

(1) For continuous variable, the range that define each factor level and the percentile of the CDF corresponding to the break point between factor levels are given. For discrete variables, each factor level is identified

(2) For this factor equal intervals are considered as levels.

For some factors in Tables 5-6 and 5-7 equal intervals are used as factor levels. The use of equal intervals for factor levels facilitates the identification of thresholds in the effect of the factor on the output of interest. For example, for the storage temperature at retail, five levels are defined with 3.5$^{o}$C increments. Using contrasts in ANOVA with this degree of level definition makes it possible to identify whether there is any temperature above which there is no temperature effect on the growth of *E. coli* organisms because of a saturation effect.

## 5.2 Analysis of Variance in the Production Module

In the production module ANOVA was applied to four parts, including feedlot prevalence, within feedlot prevalence, breeding herd prevalence, and within breeding herd prevalence. The results of the analyses are presented in Sections 5.2.1 to 5.2.4 for each of these four parts, respectively.

Table 5-8. Analysis of Variance Results for Uncertainty in the Feedlot Prevalence Part

| Variable | F Value | Pr > F | Significant | Rank |
|---|---|---|---|---|
| Study | 189,900 | <0.00001 | Yes | 1 |
| Apparent Prevalence (AP) | 1,450 | <0.00001 | Yes | 3 |
| Herd Sensitivity (HS) | 10,400 | <0.00001 | Yes | 2 |
| AP * HS | 15,000 | <0.00001 | Yes | |

### 5.2.1 Uncertainty in the Feedlot Prevalence Part

As explained in Section 3.2.1, for feedlot prevalence estimation, factors include the apparent prevalence and the herd sensitivity as quantitative factors, and the study as a qualitative one. Distributions for these factors are summarized in Table 3-9. The output in the feedlot prevalence part is the median feedlot prevalence. In Section 5.1.1, the definition of levels for each factor is explained and in Table 5-1 the assigned levels are summarized. For the feedlot prevalence part there is a one-dimensional uncertainty simulation with 65,000 iterations. Table 5-8 summarizes the result of application of ANOVA to the feedlot prevalence part.

The factors in Table 5-8 are ranked based on the magnitude of F values. Rankings are presented for statistically significant factors with Pr>F less than 0.05. Rankings are presented considering the F values only for main effects. In addition to the main effect of different factors, the interaction effect between the herd sensitivity and the apparent prevalence is also considered in the model. The F values in Table 5-8 indicate that all the factors have statistically significant effects, including the interaction term. Comparing the magnitude of F values for the main effects of different factors indicates that the study is the most sensitive factor. Hence, it is ranked first. The herd sensitivity and the apparent prevalence are ranked second and third, respectively. The difference between the F values of these three factors in the feedlot prevalence part indicates that the rankings can be considered robust. For example, the F value for the study is approximately 18 times greater than the F value for the herd sensitivity.

In order to better understand the relationship between the mean response and levels of the study, the mean response is estimated for each level of the study factor in Figure 5-5. The mean value of the feedlot prevalence is highest for the Smith (1999) study and is almost twice as large as the value estimated based upon the Hancock (1998) study. Moreover, the analyses based upon the Smith (1999) and Elder (2000) studies have approximately the same mean response.

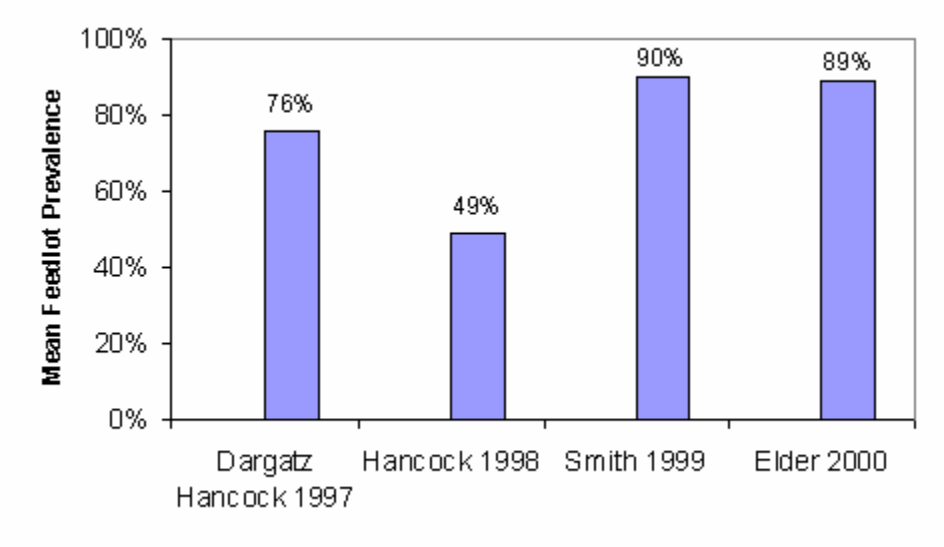



Figure 5-5. Mean Feedlot Prevalence for Levels of the Study Factor.

### 5.2.2 Uncertainty in the Within Feedlot Prevalence Part

As explained in Section 3.2.1, the quantitative factors for the within feedlot prevalence part include test sensitivity and apparent within feedlot prevalence, and the qualitative factors include study and season. Distributions for these factors are summarized in Table 3-9. Section 5.1.1 presents the definition of levels for each factor and in Table 5-2 the assigned levels are summarized. There is a one-dimensional uncertainty simulation with 65,000 iterations in this part. Table 5-9 summarizes the result of application of ANOVA to this part.

The factors in Table 5-9 are ranked based on the magnitude of F values. Rankings are presented for statistically significant factors with Pr>F less than 0.05. In addition to the main effect of different factors, the interaction effect between the test sensitivity and the apparent within feedlot prevalence, and between the study and the season are also considered. All of the factors and interaction terms have statistically significant effects. Comparing the magnitude of F values for different factors indicates that the study is the most sensitive factor. Hence, it is ranked first. The apparent within feedlot prevalence, season and test sensitivity are ranked second, third and fourth, respectively. The F value for the study is approximately 12 times greater than the F value for the second most important factor. In contrast, the F values of the second, third, and fourth most sensitive factors differ by a ratio of less than 1.3. Thus, the ranking of these factors with similar F values may not be unambiguous.

Table 5-9. Analysis of Variance Results for Uncertainty in the Within Feedlot Prevalence Part

| Variable | F Value | Pr > F | Significant | Rank |
|---|---|---|---|---|
| Study | 36,800 | <0.00001 | Yes | 1 |
| Season | 2,820 | <0.00001 | Yes | 3 |
| Apparent Within Feedlot Prevalence (AWFP) | 3,030 | <0.00001 | Yes | 2 |
| Test Sensitivity (TS) | 2,430 | <0.00001 | Yes | 4 |
| Study * Season | 4,860 | <0.00001 | Yes | |
| AWFP * TS | 80 | <0.00001 | Yes | |

There is a strong statistically significant interaction effect between the study and the season. The interaction effect between apparent within feedlot prevalence and test sensitivity is also significant, but is not as strong.

In order to visualize the relationship between the mean response and the qualitative factors, the mean response is depicted in Figure 5-6 for individual levels of these factors. Each value presented in Figure 5-6 equals the mean response at that specific factor level averaged over other factors. For example, the mean response of 22 percent for summer was estimated based on averaging over different levels of the study, test sensitivity, and apparent within feedlot prevalence when the season level equals summer. For different study levels in this figure, the smallest response value is associated with the Smith (1999) study, and the largest value is associated with the Elder (2000) study. The range of responses is from 2 to 37 percent among the different study levels. For the season factor, summer is associated with a higher value of the mean response than winter. This indicates the higher possibility of infection among feedlot cattle in summer. However, the range of the mean responses to the two levels of the season factor is not as large as the range of the mean responses for different levels of the study factor. Thus, these results confirm that the study factor is more important than the season factor.

### 5.2.3 Uncertainty in the Breeding Herd Prevalence Part

As explained in Section 3.2.1, for breeding herd prevalence estimation, factors include the apparent prevalence and the herd sensitivity as quantitative factors, and the study as a qualitative one. Distributions for these factors are summarized in Table 3-9. In Section 5.1.1, the definition of levels for each factor is explained and in Table 5-3 the assigned levels are summarized. There is a one-dimensional uncertainty simulation with 65,000 iterations in this

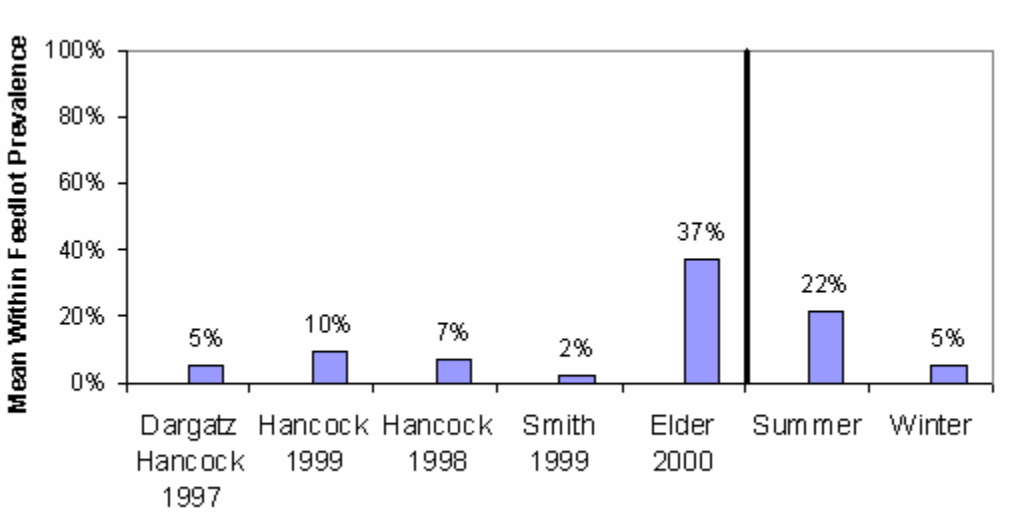



Figure 5-6. Mean Within Feedlot Prevalence for Levels of the Study and the Season Factors.

part. Table 5-10 summarizes the result of application of ANOVA to the breeding herd prevalence part.

The ranking in Table 5-10 is based on the magnitude of F values for each factor. Rankings are presented for statistically significant factors with Pr>F less than 0.05. In addition to the main effect of each factor, the interaction effect between the herd sensitivity and the apparent prevalence is also considered. F values in Table 5-10 indicate that all of the factors and the interaction term have statistically significant effects. Comparing the magnitude of F values for different factors indicates that the study is the most sensitive factor. Hence, it is ranked first. The herd sensitivity and the apparent prevalence are ranked second and third, respectively. The F value for the study is approximately 25 times greater than the F value for the second most important factor. Moreover, the F value for the herd sensitivity is approximately 9 times greater than the F value for the third most important factor.

In order to better understand the importance of the study factor, the mean response is estimated for each level of the study factor in Figure 5.7. The mean value of the breeding herd prevalence is highest for the Hancock (1998) study and is approximately 2.5 times greater than the value estimated based upon the Garber (1998) study. Moreover, the analyses based upon the Hancock (1998) and Lagreid (1998) studies have approximately the same mean response.

Table 5-10. Analysis of Variance Results for Uncertainty in the Breeding Herd Prevalence Part

| Variable | F Value | Pr > F | Significant | Rank |
|---|---|---|---|---|
| Study | 94,400 | <0.00001 | Yes | 1 |
| Apparent Prevalence (AP) | 400 | <0.00001 | Yes | 3 |
| Herd Sensitivity (HS) | 3,800 | <0.00001 | Yes | 2 |
| AP * HS | 3,760 | <0.00001 | Yes | |

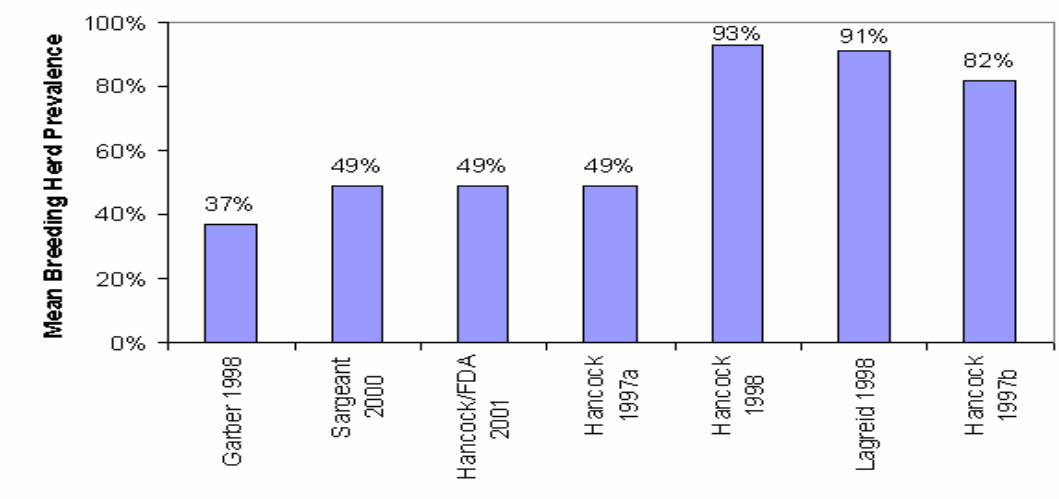



Figure 5-7. Mean Breeding Herd Prevalence for Levels of the Study Factor.

### 5.2.4 Uncertainty in the Within Breeding Herd Prevalence Part

As explained in Section 3.2.1, the quantitative factors for within breeding herd prevalence part include test sensitivity and apparent within breeding herd prevalence, and the qualitative factors include study and season. Distributions for these factors are summarized in Table 3-9. Section 5.1.1 presents the definition of levels for each factor and in Table 5-4 the assigned levels are summarized. There is a one-dimensional uncertainty simulation with 65,000 iterations in this part. Table 5-11 summarizes the result of application of ANOVA to this part.

In addition to the main effect of each factor, the interaction effects between test sensitivity and apparent within breeding herd prevalence, and between study and season are also considered. All of the individual factors are statistically significant, while none of the interaction terms are statistically significant interaction. Study is the most sensitive factor. Apparent within breeding herd prevalence, test sensitivity and season are ranked second, third and fourth, respectively. The F value for is approximately 9 times greater than the F value for the second

Table 5-11. Analysis of Variance Results for Uncertainty in the Within Breeding Herd Prevalence Part

| Variable | F Value | Pr > F | Significant | Rank |
|---|---|---|---|---|
| Study | 18,830 | <0.00001 | Yes | 1 |
| Season | 12 | 0.0005 | Yes | 4 |
| Apparent Within Breeding Herd Prevalence (AWBHP) | 2,100 | <0.00001 | Yes | 2 |
| Test Sensitivity (TS) | 400 | <0.00001 | Yes | 3 |
| Study * Season | 2 | 0.10 | No | |
| AWBHP * TS | 1 | 0.4 | No | |

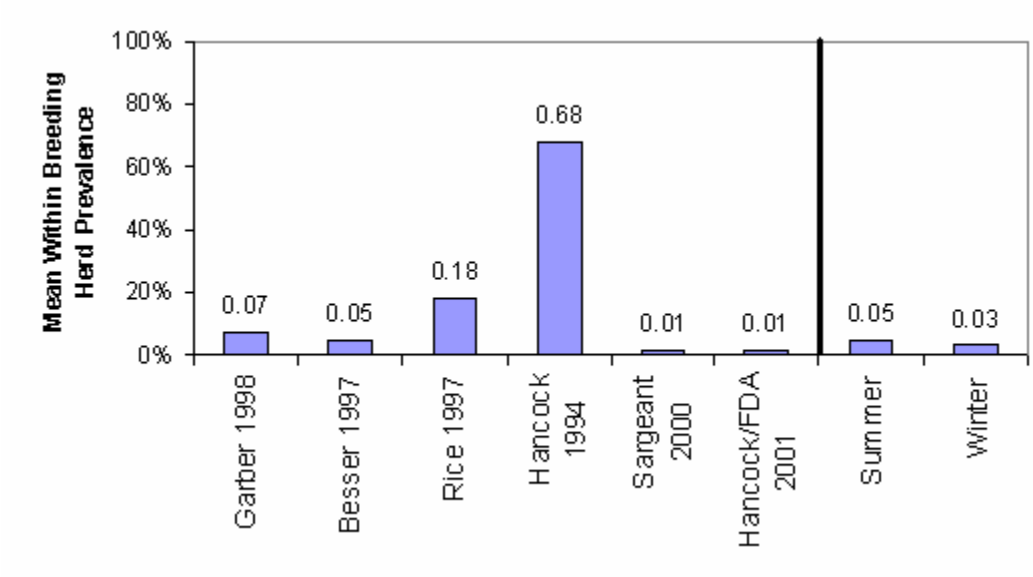



Figure 5-8. Mean Within Breeding Herd Prevalence for Levels of the Study and the Season Factors.

most important factor. The relative differences of F values among other factors are also large. Therefore, the rankings are considered to be unambiguous.

In order to visualize the relationship between the mean response and qualitative factors (i.e., the study and the season) the mean response is depicted in Figure 5-8 for individual levels of these factors. Each value presented in Figure 5-8 equals the mean response at that specific factor level averaged over other factors. For example, the mean response of 5 percent for summer was estimated based on averaging over different levels of the study, test sensitivity, and apparent within breeding herd prevalence for the summer season. For different study levels in this figure, the smallest response value is associated with the Sargeant (2000) and Hancock/CFSAN (2001) studies, and the largest value is associated with the Hancock (1994) study. The range of responses is from 1 to 68 percent among the different study levels. The

infection prevalence value in summer differs from the infection prevalence in winter by a ratio of only 1.6 indicating that for breeding herds seasonal effect does not affect the infection prevalence substantially. Moreover, the range of the mean responses to the two levels of the season factor is not as large as the range of the mean responses for different levels of the study factor. Thus, these results confirm that the study factor is more important than the season factor.

## 5.3 Analysis of Variance in the Slaughter Module

The slaughter module is discussed in Section 3.2.2. Factors and corresponding distributions in the slaughter module are summarized in Table 3-10. Three different types of probabilistic analysis were performed for this module, as described in Section 3.3.2: (1) one-dimensional simulation of variability based upon mean values of uncertain factors; (2) two-dimensional simulation of variability for each realization of uncertainty; and (3) one-dimensional simulation of both variability and uncertainty co-mingled. In this section, the results of ANOVA for each of these three types of simulations are given. The case study scenario for the slaughter module is focused upon steers and heifers in the high prevalence season. Section 5.1.2 presents the definition of levels for each factor and in Table 5-5 the assigned levels are summarized.

In the next section, the results of ANOVA are presented based upon simulation of variability only. In Section 5.3.2, results are presented based upon the two-dimensional simulation of variability for different realizations of uncertainty. Results for the co-mingled one-dimensional simulation of both variability and uncertainty are given in Section 5.3.3. Section 5.3.4 compares the results from Sections 5.3.1 to 5.3.3.

### 5.3.1 Variability Only

This section presents the results of ANOVA applied to a one dimensional probabilistic simulation in which variability is only considered for mean value of uncertain factors, based upon the case study scenario described in Section 3.3.2. The factor levels used in this analysis are the same as those given in Table 5-5.

Table 5-12 summarizes the results of application of ANOVA to the slaughter module for the simulation of variability only. Rankings are presented for statistically significant factors with Pr>F less than 0.05. In addition to the main effect of each factor, interaction effects are also considered between: (1) the chilling effect and the Trim/Vacuum/Washing efficiency; (2) the number of organisms and the Trim/Vacuum/Washing efficiency; (3) the number of organisms

Table 5-12. The Analysis of Variance Results for Steer and Heifer Combo Bin Contamination in Summer Based Upon Variability only

| Variable | F Value | Pr > F | Significant | Rank |
|---|---|---|---|---|
| Total Number of Combo Bins for Each Carcass (TNCB) | 2.1 | 0.08 | No | --- |
| Total Number of Infected Animals (TNI) | 0.1 | 0.9 | No | --- |
| Total Number of Contaminated Animals (TNC) | 7.2 | 0.0007 | Yes | 9 |
| Probability of Positive Cases at Both Steps of Dehiding and Evisceration ($P_{both}$) | 0.4 | 0.8 | No | --- |
| Number of Positive Cases at Both Steps of Dehiding and Evisceration ($N_{both}$) | 282 | <0.0001 | Yes | 5 |
| Number of Positive Cases at Evisceration (NPE) | 33 | <0.0001 | Yes | 8 |
| Chilling Effect ($CH_{eff}$) | 1480 | <0.0001 | Yes | 1 |
| Number of Organisms ($N_{org}$) | 850 | <0.0001 | Yes | 3 |
| Trim/Vacuum/Washing Efficiency (TVW) | 1030 | <0.0001 | Yes | 2 |
| Evisceration Organisms Added ($N_{evisc}$) | 143 | <0.0001 | Yes | 6 |
| Washing Effect ($W_{eff}$) | 492 | <0.0001 | Yes | 4 |
| Contaminated $cm^2$ (CCM) | 50 | <0.0001 | Yes | 7 |
| $CH_{eff}$ * TVW | 487 | <0.0001 | Yes | |
| $N_{org}$ * TVW | 28 | <0.0001 | Yes | |
| $N_{org}$ * $W_{eff}$ | 815 | <0.0001 | Yes | |
| $N_{evisc}$ * $W_{eff}$ | 304 | <0.0001 | Yes | |

and the washing effect; and (4) the evisceration organisms added and the washing effect are considered. Three factors are not statistically significant, including the total number of combo bins to which each carcass contributes, the total number of infected animals, and the probability of positive cases at both steps of dehiding and evisceration. All four interaction terms are statistically significant.

Comparing the magnitude of F values for the statistically significant factors indicates that the chilling effect, Trim/Vacuum/Washing efficiency, and the number of organisms are the three most sensitive factors. The relative difference of F values for these factors indicates that these three factors may be of comparable importance. For example, the F values for the chilling effect and the Trim/Vacuum/Washing efficiency differ by a ratio of only 1.4. The ambiguity of these rankings is further evaluated in Section 5.3.2, when the factors are ranked for different uncertainty realizations.

For factors rather than the three most sensitive ones, the F values are comparatively small, especially for the seventh through ninth ranked inputs. Although all interactions have statistically significant effects, the interactions between the number of organisms on contaminated carcass and the washing effect, and between the chilling effect and the Trim/Vacuum/Washing efficiency, are more important than others considered based upon the magnitude of their F values.

### 5.3.2 Two-Dimensional Simulation of Variability for Different Uncertainty Realizations

The application of ANOVA to a two-dimensional simulation in which variability is simulated for each different realization of uncertainty involves sensitivity analysis for each of the uncertainty iterations. In this case, for example, there are 100 uncertainty iterations. Within each uncertainty iteration, 650 samples were generated to represent variability in each factor. Thus, ANOVA was applied 100 times. The factor levels used in this analysis are the same as those given in Table 5-5.

The factors included in ANOVA for the two-dimensional simulation were the same as those for the one-dimensional simulation of variability only as listed in Table 5-12 without considering the interaction terms. The interaction terms were not considered because the datasets were unbalanced. This means that the sample sizes within levels were different, with some levels having comparatively small sample sizes. For some uncertainty realizations the lack of balance is severe enough to lead to singularities in the solution algorithm.

The results of the 100 analyses with ANOVA are summarized in Table 5-13. The table includes the mean F value of the factor and the minimum to maximum range of F values over the 100 simulations. The percentage of the 100 simulations that produced a statistically significant F value is quantified. Furthermore, the mean rank and the range of ranks are given for each factor.

The mean ranks indicate that the chilling effect is the most important factor. There is 100 percent probability that the chilling effect is identified as a statistically significant factor among all 100 uncertainty realizations. The mean ranks for the Trim/Vacuum/Washing efficiency, washing efficiency, and the number of organisms are estimated as 4.2, 4.4, and 4.4 indicating that on average the output has approximately similar sensitivity to these factors. For these factors the probabilities of being statistically significant in 100 uncertainty realizations are 78, 75, and 74 percent, respectively. However, although these factors have approximately similar average

Table 5-13. Summary of the ANOVA Results for Two-Dimensional Variability Simulation for 100 Uncertainty Realizations

| Variable | Mean F Value | Minimum F Value | Maximum F Value | Frequency (1) | Mean Rank | Range of Rank |
|---|---|---|---|---|---|---|
| TNCB | 1.3 | 0.06 | 15.2 | 12 | 9.6 | 2 – 12 |
| TNI | 1.4 | 0.01 | 0.8 | 17 | 9.4 | 3 – 12 |
| TNC | 752 | 0.01 | 21,700 | 69 | 5.8 | 1 - 12 |
| $P_{both}$ | 1.3 | 0.02 | 4.2 | 6 | 9.1 | 3 – 12 |
| $N_{both}$ | 89 | 0.01 | 4,605 | 52 | 7.5 | 2 – 12 |
| NPE | 449 | 0.01 | 15,560 | 48 | 7.4 | 1 – 12 |
| $CH_{eff}$ | 128,000 | 3.73 | 3,490,000 | 100 | 1.7 | 1 – 6 |
| $N_{org}$ | 9,200 | 0.01 | 442,550 | 74 | 4.4 | 1 – 12 |
| TVW | 2,700 | 0.1 | 105,000 | 78 | 4.2 | 1 – 11 |
| $N_{evisc}$ | 5,900 | 0.01 | 503,000 | 37 | 8.0 | 1 – 12 |
| $W_{eff}$ | 12,100 | 0.02 | 884.600 | 75 | 4.4 | 1 – 11 |
| CCM | 1,100 | 0.01 | 28,860 | 65 | 5.8 | 1 - 12 |

(1) The percentage of the 100 uncertainty simulations for which the F value was statistically significant.

and range of rankings indicating that they are of comparable importance to each other, they are less important than the chilling effect. The number of contaminated $cm^2$ of meat trims and the total number of contaminated animals each have a mean rank of 5.8 with a probability of being statistically significant of 65 and 69 percent, respectively. Thus, the output has approximately similar sensitivity to these two factors. Moreover, the output has approximately similar sensitivity to the number of positive cases at evisceration and the number of positive cases at both steps of dehiding and evisceration with mean ranks of 7.4 and 7.5, respectively, in 100 uncertainty realizations. The output has the lowest sensitivity to the number of combo bins to which each animal contributes, with mean rank of 9.6. This factor was statistically significant in only 12 percent of the uncertainty realizations.

In order to visualize the results of the sensitivity analysis, the complementary cumulative distribution function (CCDF) of the rank is given for each factor based upon the 100 uncertainty realizations in Figures 5-9 to 5-11. Figure 5-9 displays the CCDFs for four factors that have the highest average ranks among all of the factors included in the analysis. These factors are chilling effect ($CH_{eff}$), washing efficiency ($W_{eff}$), number of organisms on the carcass surface ($N_{org}$), and Trim/Vacuum/Washing efficiency (TVW). The CCDF for the chilling effect indicates that for approximately 52 percent of the simulations, the rank was one, which implies that the rank was worse than one for 48 percent of the simulations. Furthermore, the chilling effect was ranked

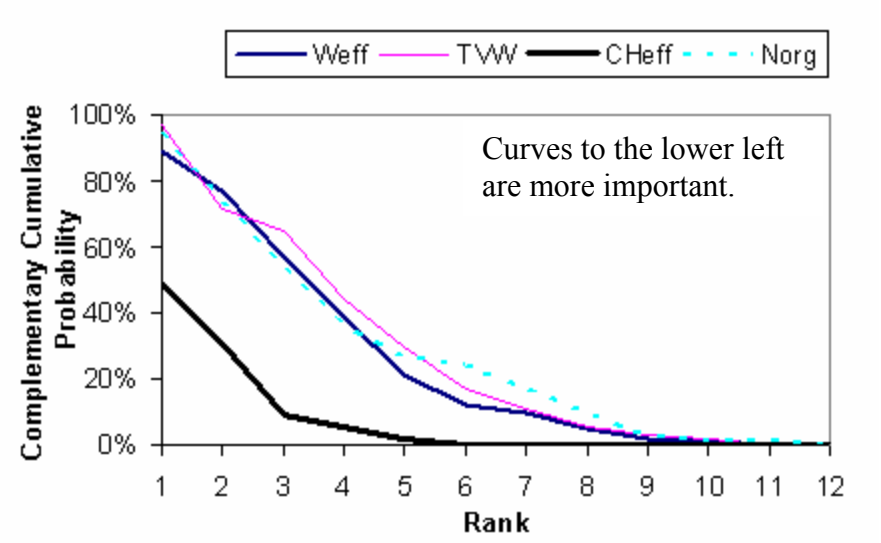



Figure 5-9. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected factors: Chilling Effect ($CH_{eff}$); Washing Efficiency ($W_{eff}$); Trim/Vacuum/Washing Efficiency (TVW); and Number of Organisms on the Carcass Surface ($N_{org}$).

five or higher for 96 percent of the simulations. In contrast, washing efficiency was ranked first for 12 percent of the simulations and was ranked fifth or higher for 80 percent of the simulations. The distributions of ranks for washing efficiency, Trim/Vacuum/Washing efficiency, and number of organisms on contaminated carcasses are similar to each other. Thus, although the chilling effect has the highest frequency of a rank of one, there is some ambiguity regarding which of the other three factors factor is the second most important.

When comparing the CCDFs of Figure 5-9, it is apparent that the chilling effect tends to have a higher rank than the other factors. Furthermore, because the probability that the chilling effect has a rank of five or higher is nearly 100 percent, the identification of the chilling effect as one of the most important factors is robust to uncertainty. In contrast, the washing efficiency, Trim/Vacuum/Washing efficiency, and the number of organisms have 20, 28, and 30 percent probability, respectively, of having a rank worse than five. Thus, although these three factors typically have a similar importance to each other, they are typically less important than the chilling effect.

Figure 5-10 displays the CCDFs for five factors that have the highest probability of a middle range of average ranks between five and eight among all of the factors included in the analysis. These factors are total number of contaminated animals (TNC), contaminated $cm^2$ of

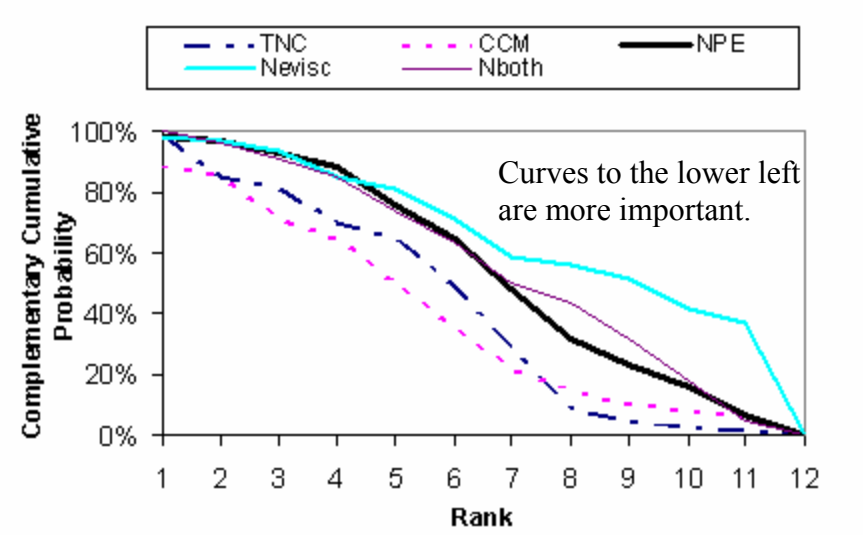



Figure 5-10. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected factors: Total Number of Contaminated Animals (TNC); Contaminated $Cm^2$ of Meat Trims (CCM); Number of Positive Cases at Evisceration (NPE); Number of *E. coli* Organisms Added Due to Evisceration ($N_{evisc}$); and Number of Positive Cases at both Steps of Dehiding and Evisceration ($N_{both}$).

meat trims (CCM), number of positive cases at evisceration (NPE), number of organisms added due to evisceration ($N_{evisc}$), and number of positive cases at both steps of dehiding and evisceration ($N_{both}$). The CCDF for the contaminated $cm^2$ of meat trims indicates that for approximately 88 percent of the simulations, the rank was worse than one, which implies that the rank was equal to one for only 12 percent of the simulations. In contrast, for other factors, the probability of the rank being worse than one is 98 percent.

Although the mean ranks of these factors vary between 6 and 8, the probability that the ranks are worse than eight varies from 10 to 58 percent among the five selected factors. Furthermore, these five factors have ranks ranging from as high as one to as low as 12 in some cases. Thus, it is apparent that the identification of the rank of these factors is not robust to uncertainty. Hence, there is ambiguity regarding the rank of each factor as a function of uncertainty in the model factors.

The least important group of factors is depicted in Figure 5-11. These factors include the total number of combo bins to which each animal contributes (TNCB), probability of positive cases at both steps of dehiding and evisceration ($P_{both}$), and the total number of infected animals (TNI). These factors have a probability ranging from 93 to 98 percent of having a rank worse than five, and their average ranks range from 9 to 10. These factors have similar CCDF

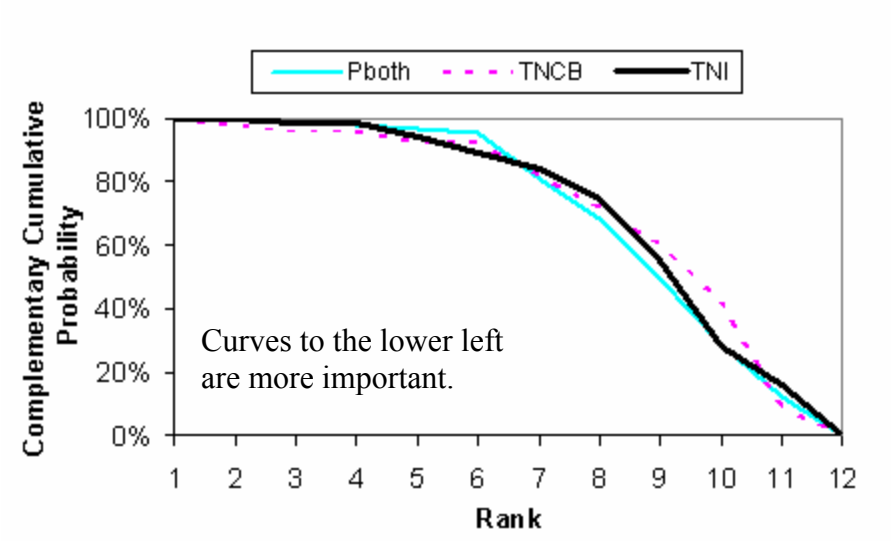



Figure 5-11. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected factors: Probability of Positive Cases at both Steps of Dehiding and Evisceration ($P_{both}$); Total Number of Combo Bins (TNCB); and Total Number of Infected Animals (TNI).

distributions. The similarity of these distributions implies that these three factors are of comparable importance. However, even though all three of these factors are typically ranked seven or worse for approximately 80 percent of the uncertainty realizations, there are a few uncertainty iterations for which these factors have ranks as high as two or three. Hence, there is ambiguity regarding the rank of each factor as a function of uncertainty in the model factors, although with high probability the ranks are worse than six in different uncertainty realizations. Furthermore, even taking into account uncertainty, these three factors are clearly less important than the most important input, chilling effect.

### 5.3.3 One-Dimensional Simulation of Variability and Uncertainty

This section presents the results of ANOVA applied to a one dimensional probabilistic simulation in which variability and uncertainty are co-mingled, based upon the case study scenario described in Section 3.3.2. The factor levels used in this analysis are the same as those given in Table 5-5.

Table 5-14 summarizes the results of application of ANOVA to the slaughter module for the co-mingled simulation of variability and uncertainty. The factors in Table 5-14 are ranked based on the magnitude of F values. Rankings are presented for statistically significant factors with Pr>F less than 0.05. In addition to the main effect of each factor, the interaction effects

Table 5-14. The Analysis of Variance Results for Steer and Heifer Combo Bin Contamination in Summer Based Upon One-Dimensional Co-Mingled Variability and Uncertainty Simulation

| Variable | F_Value | Pr > F | Significant | Rank |
|---|---|---|---|---|
| Total Number of Combo Bins for Each Carcass | 0.5 | 0.7 | No | --- |
| Total Number of Infected Animals | 0.7 | 0.5 | No | --- |
| Total Number of Contaminated Animals | 12 | <0.0001 | Yes | 6 |
| Probability of Positive Cases at both Steps of Dehiding and Evisceration | 1.2 | 0.3 | No | --- |
| Number of Positive Cases at both Steps of Dehiding and Evisceration | 2.7 | 0.04 | Yes | 8 |
| Number of Positive Cases at Evisceration | 7.8 | <0.0001 | Yes | 7 |
| Chilling Effect ($CH_{eff}$) | 1053 | <0.0001 | Yes | 1 |
| Number of Organisms ($N_{org}$) | 253 | <0.0001 | Yes | 2 |
| Trim/Vacuum/Washing Efficiency (TVW) | 225 | <0.0001 | Yes | 3 |
| Evisceration Organisms Added ($N_{evisc}$) | 76 | <0.0001 | Yes | 5 |
| Washing Effect ($W_{eff}$) | 159 | <0.0001 | Yes | 4 |
| Contaminated $cm^2$ | 2.7 | 0.04 | Yes | 8 |
| $CH_{eff}$ * TVW | 95 | <0.0001 | Yes | |
| $N_{org}$ * TVW | 16 | <0.0001 | Yes | |
| $N_{org}$ * $W_{eff}$ | 124 | <0.0001 | Yes | |
| $N_{evisc}$ * $W_{eff}$ | 60 | <0.0001 | Yes | |

between the chilling effect and the Trim/Vacuum/Washing efficiency, between the number of organisms and the Trim/Vacuum/Washing efficiency, between the number of organisms and the washing effect, and between the evisceration organisms added and the washing effect are considered. F values in Table 5-14 indicate that there are no statistically significant effects for factors such as total number of combo bins to which each carcass contributes, the total number of infected animals, and the probability of positive cases at both steps of dehiding and evisceration. Moreover, F values indicate that the interaction terms have statistically significant effects.

Comparing the magnitude of F values for the statistically significant factors indicates that the chilling effect, the number of organisms, and Trim/Vacuum/Washing efficiency are the three most sensitive factors. The relative difference of F values for these factors indicates that the rank of the most sensitive factor is substantially different, although the rank of the second and third important factors may be comparable. The F values for the chilling effect and the number of organisms differ by a ratio of 4.2. In contrast, the F values for the number of organisms and Trim/Vacuum/Washing efficiency differ by a ratio of only 1.1. The robustness of these rankings

Table 5-15. Evaluation of ANOVA Contrasts Regarding the Interactions Between the Chilling Effect and the Trim/Vacuum/Washing Efficiency

| Contrast | Estimate | F Value | Pr>F | Significant |
|---|---|---|---|---|
| Comparing $TVW_{eff}>75^{\%}$ and $TVW_{eff}<60^{\%}$ when there is 1log increase in number of organisms due to chilling | -2.2 | 0.2 | 0.7 | No |
| Comparing $TVW_{eff}>75^{\%}$ and $TVW_{eff}<60^{\%}$ when there is 2logs increase in number of organisms due to chilling | -3.6 | 0.2 | 0.7 | No |
| Comparing $TVW_{eff}>75^{\%}$ and $TVW_{eff}<60^{\%}$ when there is more than 2logs increase in number of organisms due to chilling | 442 | 430 | <0.0001 | Yes |

was evaluated in Section 5.3.2, when the factors were ranked for different uncertainty realizations.

For factors other than the three most sensitive ones, small F values indicate that these factors may be unimportant. For example, F values for the evisceration organisms added, the total numbers of contaminated animals, the number of positive cases at evisceration, and the contaminated $cm^2$ of meat trims, which are ranked fifth to eighth, differ by ratios of 3 to 83 compared to the F value of the third important factor.

Although all interactions have statistically significant effects, the interactions between the number of organisms on contaminated carcass and the washing effect, and between the chilling effect and the Trim/Vacuum/Washing efficiency are the most important based upon the magnitude of F values.

In addition to the inferences obtained by application of ANOVA regarding the sensitivity of the output to individual factors, additional information regarding sensitivity is achieved by using contrasts. Contrasts are useful in order to find thresholds in the model response to different factors, or in understanding the response of the model to interactions between factors. As discussed in Section 2.2.2, not all contrasts are estimable in unbalanced experiments, because it is possible that there are not enough data to estimate the result of a contrast.

Contrasts were prepared only for the mixed analysis, since this analysis takes into account the widest range of possible values of each factor in the context of a single simulation. Specification of contrasts in ANOVA is done manually and it cannot be automated. Hence, it was not practical to perform contrasts for the two-dimensional analysis, which would have

required repeating manual analysis 100 times. The selection of factors to include in the contrasts was based upon the sensitivity results of the individual factors. The rankings in Table 5-14 indicate that the chilling effect is the most important factor. In addition, there is a significant interaction effect between the chilling effect and the Trim/Vacuum/Washing efficiency. Hence, a set of contrasts was prepared to evaluate the response of the model to the interaction between these two factors. The results of these contrasts are summarized in Table 5-15.

The contrasts in Table 5-15 compare the mean response in the slaughter module at different levels of the chilling effect when the Trim/Vacuum/Washing efficiency varies between the highest level (e.g., $TVW_{eff}<60^{\%}$) and the lowest level (e.g., $TVW_{eff}>75^{\%}$). The 'Estimate' column in Table 5-15 presents the estimate of the difference between the mean responses for the condition mentioned in the contrast. If the estimate is not significant, which means that the Pr>f is greater than 0.05, there is not enough statistical support indicating that the estimated value for the contrast is different from zero.

Results in Table 5-15 indicate that changing the efficiency of the decontamination step from low to high can only affect the contamination in combo bins when there is more than 2 logs increase occurred in the number of *E. coli* organisms during the chilling process. Otherwise, if the amount of *E. coli* organisms on carcasses does not increase more than 2 logs during the chilling process, there is no statistically significant difference in the final combo bin contamination when applying different efficiencies in the decontamination step (i.e., Trim/Vacuum/Washing step). For example, the first contrast in Table 5-15 indicates that when there is less than 1 log increase in the number of *E. coli* organisms on carcasses during the chilling process, there is no statistically significant difference between the combo bin contaminations when applying high and low efficiency of decontamination by using the Trim/Vacuum/Washing step. The F value of 430 in the last contrast indicates that for carcasses that had an increase of more than 2 logs in the number of *E. coli* organisms on their surfaces during the chilling process, there are on average 442 more *E. coli* organisms in the combo bins filled with meat trims coming from these carcasses, when using the low level of efficiency during the decontamination step.

Table 5-16. Summary of the ANOVA Results in the Slaughter Module Based on Variability Only, Variability for Different Uncertainty Realizations, and Co-mingled Variability and Uncertainty Analyses

| Variable | Ranks | | |
|---|---|---|---|
| | Analysis 1(1) | Analysis 2(2) | Analysis 3(3) |
| Total Number of Combo Bins for Each Carcass | --- | 9.6 | --- |
| Total Number of Infected Animals | --- | 9.4 | --- |
| Total Number of Contaminated Animals | 9 | 5.8 | 6 |
| Probability of Positive Cases at Both Steps of Dehiding & Evisceration | --- | 9.1 | --- |
| Number of Positive Cases at Both Steps of Dehiding & Evisceration | 5 | 7.5 | 8 |
| Number of Positive Cases at Evisceration | 8 | 7.4 | 7 |
| Chilling Effect | 1 | 1.7 | 1 |
| Number of Organisms | 3 | 4.4 | 2 |
| Trim/Vacuum/Washing Efficiency | 2 | 4.2 | 3 |
| Evisceration Organisms Added | 6 | 8.0 | 5 |
| Washing Effect | 4 | 4.4 | 4 |
| Contaminated $cm^2$ | 7 | 5.8 | 8 |

(1) Ranks based on the variability only analysis.
(2) Mean ranks based on the variability for different uncertainty realizations analysis.
(3) Ranks based on the one-dimensional co-mingled variability and uncertainty analysis.

### 5.3.4 Summary and Comparison of the Results of ANOVA in the Slaughter Module

In Sections 5.3.1 to 5.3.3 ANOVA was applied to three datasets considering variability only, variability for different uncertainty realizations, and co-mingled variability and uncertainty in factors. In this section rankings based on these analyses are summarized and compared. Table 5-16 gives the ranks for each factor based on analyses in Sections 5.31 to 5.3.3.

Table 5-16 indicates that the chilling effect is identified as the most important factor based upon all three simulations. The number of organisms was identified as the second most important factor based upon the co-mingling of variability and uncertainty and as the third most important factor based upon variability only. However, for both of these simulations, the F value for this factor was not substantially different than the F value for the Trim/Vacuum/Wash efficiency. Furthermore, the group of top three factors was the same for both simulation methods. Moreover, in the second analysis taking into account the uncertainty in factors, the Trim/Vacuum/Wash efficiency and the number of organisms and have mean ranks of 4.2 and

4.4. Since there is no factor identified with higher mean rank in this analysis, this indicates that these two factors are placed after the chilling effect regarding their importance. Hence, all three simulations have agreement on the top four sensitive factors. The variability only and the co-mingled variability and uncertainty analyses identified the total number of combo bins to which animals contributes, the total number of infected animals, and the probability of positive cases at both steps of dehiding and evisceration as statistically insignificant factors. These factors have mean ranks of 9.6, 9.4, and 9.1 with the second analysis with frequency of being significant of 13.6, 19.3, and 6.8 percent, respectively, in 100 uncertainty iterations.

Therefore, the key similarities among the three probabilistic simulations were with respect to the identification of the most important factor, secondary importance factors, and the least important factors. Factors that were of moderate importance based upon each of three methods were not completely similar. For example, the number of positive cases at both steps of dehiding and evisceration has a rank of 5 with variability only analysis, while it has mean rank of 7.5 and rank of 8 with the second and third analyses, respectively.

## 5.4 Analysis of Variance in the Preparation Module

In the preparation module ANOVA was applied to three parts, including growth estimation, cooking effect, and serving contamination parts. The results of the analyses are presented in Sections 5.4.1, 5.4.3, and 5.4.4 for each of these three parts. Moreover, Section 5.4.2 presents a discussion regarding the sampling distribution of F values.

### 5.4.1 Analysis of Variance in the Growth Estimation Part

The growth estimation part is discussed in Section 3.2.3. Three different types of probabilistic analysis were performed for this part, as described in Section 3.3.3: (1) one-dimensional simulation of variability based upon mean values of uncertain factors; (2) two-dimensional simulation of variability for each realization of uncertainty; and (3) one-dimensional simulation of both variability and uncertainty co-mingled. Section 5.1.3 presents the definition of levels for each factor and in Table 5-7 the assigned levels are summarized. In this section, the results of ANOVA for each of these three types of simulations are given.

In the next section, the results of ANOVA are presented based upon simulation of variability only. In Section 5.4.1.2, results are presented based upon the two-dimensional simulation of variability for different realizations of uncertainty. Results for the co-mingled one-dimensional simulation of both variability and uncertainty are given in Section 5.4.1.3. Section

Table 5-17. The Analysis of Variance Results for the Growth Estimation Part Based Upon Variability only ($R^2 = 0.81$)

| Variable | F Value | Pr > F | Significant | Rank |
|---|---|---|---|---|
| Storage Temperature, Stage 1 ($Temp_1$) | 940 | <0.0001 | Yes | 3 |
| Storage Temperature, Stage 2 ($Temp_2$) | 3.7 | 0.03 | Yes | 10 |
| Storage Temperature, Stage 3 ($Temp_3$) | 1,240 | <0.0001 | Yes | 2 |
| Storage Time, Stage 1 ($Time_1$) | 930 | <0.0001 | Yes | 4 |
| Storage Time, Stage 2 ($Time_2$) | 0.6 | 0.4 | No | --- |
| Storage Time, Stage 3 ($Time_3$) | 5,390 | <0.0001 | Yes | 1 |
| Maximum Density (MD) | 25 | <0.0001 | Yes | 9 |
| Lag Period, Stage 1 ($LP_1$) | 300 | <0.0001 | Yes | 5 |
| Lag Period, Stage 2 ($LP_2$) | 3 | 0.03 | Yes | 11 |
| Lag Period, Stage 3 ($LP_3$) | 230 | <0.001 | Yes | 6 |
| Generation Time, Stage 1 ($GT_1$) | 40 | <0.0001 | Yes | 8 |
| Generation Time, Stage 2 ($GT_2$) | 0.4 | 0.7 | No | --- |
| Generation Time, Stage 3 ($GT_3$) | 45 | <0.0001 | Yes | 7 |
| $Temp_1$ * $Time_1$ | 1,340 | <0.0001 | Yes | |
| $Temp_2$* $Time_2$ | 0.3 | 0.6 | No | |
| $Temp_3$ * $Time_3$ | 3,400 | <0.0001 | Yes | |

5.4.1.4 compares the results from Sections 5.4.1.1 to 5.4.1.3.

As an example case study, the coefficient of determination, $R^2$, is provided for the three types of probabilistic analysis performed in this part. This coefficient represents the amount of output variation captured by the model considering the main effects of the factors and the interaction effects between selected factors. Low values of $R^2$ indicate that there may be additional terms that should be included in the model. Those terms could capture a higher amount of variation in the output. Such additional terms may include main effect of other factors or higher order interaction terms, such as three or four way interactions.

5.4.1.1 Variability Only

This section presents the results of ANOVA applied to a one-dimensional probabilistic simulation in which variability is only considered for mean uncertainty, based upon the case study scenario described in Section 3.3.3. The factor levels used in this analysis are the same as those given in Table 5-7.

Table 5-17 summarizes the results of application of ANOVA to the growth estimation part for the simulation of variability only. The factors in Table 5-17 are ranked based on the magnitude of F values. Rankings are presented for statistically significant factors with Pr>F less

than 0.05. In addition to the main effect of each factor, the interaction effects between the storage time and the storage temperature at stages 1 to 3 are considered.

F values in Table 5-17 indicate that there are no statistically significant effects for factors such as the storage time and the generation time at stage 2. Moreover, The F values indicate that there is no statistically significant interaction between the storage time and the storage temperature at stage 2. Comparison of the F values for the interaction terms indicates that the interaction between the storage time and the storage temperature at stage 3 has higher importance that the interaction between these two factors at stage 1. The F value for the interaction effect between the storage temperature and storage time at stage 3 differs from the F value for the interaction between these two factors in stage 1 by ratio of 2.6.

Comparing the magnitude of F values for the statistically significant factors indicates that the storage time at stage 3, the storage temperature at stage 3, the storage time at stage 1, and the storage temperature at stage 1 are the four most sensitive factors. The relative difference of F values for these factors indicates that the ranking of the top factor is robust. The F values for the storage time and the storage temperature at stage 3 differ by a ratio of 4.4 indicating that the rank of the storage time at stage 3 is robust. In contrast, comparison of the F values for the storage temperature at stage 3 and the storage time at stage 1 indicates that the relative ranking of these two factors are not robust. The F values for these factors differ by a ratio of only 1.3. The robustness of these rankings is further evaluated in Section 5.4.1.2, when the factors are ranked for different uncertainty realizations.

For factors rather than the four most sensitive ones, small F values indicate that these factors may be unimportant. For example, F values for the generation time at stage 1, the maximum density, the storage temperature at stage 2, and the lag period at stage 2 which are ranked eight to eleven, differ by ratios of 23, to 1800 compared to the F value of the fourth important factor.

The coefficient of determination, $R^2$, is 0.81. This high value of $R^2$ indicates that a substantial portion of the variation in the output is attributable to the effects included in the analysis, including both the main effects of factors and two way interactions between selected factors. Although it may be possible to increase the coefficient of determination by including additional effects in the analysis, such as additional two way interactions or by including higher

Table 5-18. Summary of the ANOVA Results for Two-Dimensional Variability Simulation for Different Uncertainty Realizations (Mean $R^2 = 0.63$)

| Variable | Mean F Value | Minimum F Value | Maximum F Value | Frequency (1) | Mean Rank | Range of Rank |
|---|---|---|---|---|---|---|
| Temp1 | 45.2 | 0.3 | 1,208 | 86 | 4.0 | 1 – 11 |
| Temp2 | 1.0 | 0.0 | 6 | 5 | 9.6 | 4 – 13 |
| Temp3 | 72.3 | 10.6 | 580 | 100 | 2.5 | 1 – 4 |
| Time1 | 45.1 | 0.8 | 406 | 99 | 2.7 | 1 – 6 |
| Time2 | 0.3 | 0.0 | 5 | 1 | 12.1 | 6 – 13 |
| Time3 | 143.8 | 28.1 | 1,250 | 100 | 1.3 | 1 – 4 |
| MD | 1.1 | 0.1 | 9 | 7 | 9.7 | 5 – 13 |
| LP1 | 3.9 | 0.6 | 22 | 69 | 6.4 | 4 – 11 |
| LP2 | 1.1 | 0.0 | 7 | 13 | 9.2 | 5 – 13 |
| LP3 | 3.9 | 0.1 | 15 | 73 | 6 | 3 – 11 |
| GT1 | 1.4 | 0.3 | 22 | 10 | 9.3 | 4 – 13 |
| GT2 | 1.1 | 0.0 | 6 | 9 | 9.3 | 5 – 13 |
| GT3 | 1.5 | 0.2 | 12 | 10 | 8.8 | 5 - 13 |

(1) The percentage of the 100 uncertainty simulations for which the F value was statistically significant.

#### 5.4.1.2 Two-Dimensional Simulation of Variability for Different Uncertainty Realizations

order interactions, the coefficient of determination based upon this analysis is sufficiently high to confirm that most of the variation in the output is accounted for. Therefore, in this case, no additional refinement was made to the analysis.

The application of ANOVA to a two-dimensional simulation in which variability is simulated for each different realization of uncertainty involves sensitivity analysis for each of the uncertainty iterations. In this case, for example, there are 100 uncertainty iterations. Within each uncertainty iteration, 650 samples were generated to represent variability in each factor. Thus, ANOVA was applied 100 times. The factor levels used in this analysis are the same as those given in Table 5-7.

The factors included in ANOVA for the two-dimensional simulation were the same as those for the one-dimensional simulation of variability only as listed in Table 5-17 without considering the interaction terms.

The results of the 100 analyses with ANOVA are summarized in Table 5-18. The table includes the mean F value and the minimum to maximum range of F values over the 100 simulations. The percentage of the 100 simulations that produced a statistically significant F

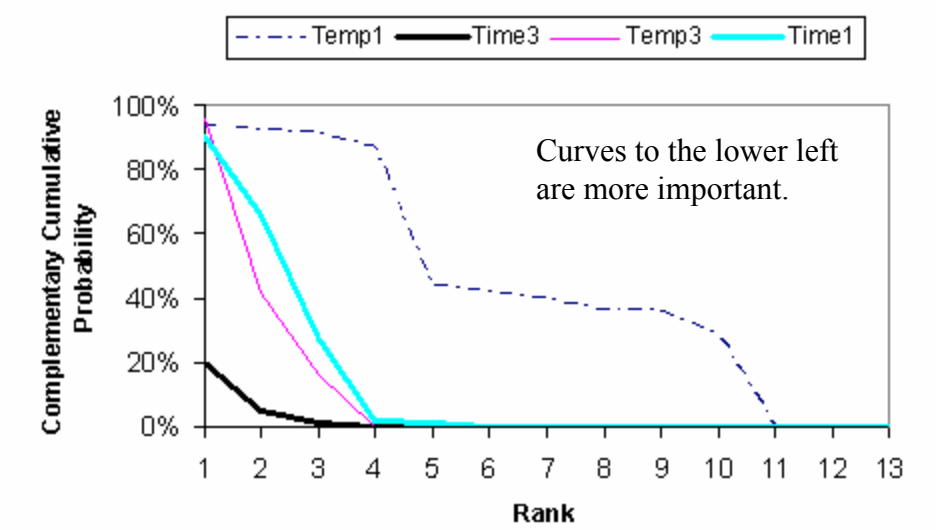



Figure 5-12. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected Factors: Storage Temperature at Stages 1and 3 ($Temp_1$ and $Temp_3$); and Storage Time at Stages 1 and 3 ($Time_1$ and $Time_3$).

value is quantified. Furthermore, the mean rank and the range of ranks are given for each factor. The mean ranks indicate that the storage time at stage 3 is the most important factor. There is 100 percent probability that this factor is identified as statistically significant in the uncertainty realizations. The mean ranks for the storage temperature at stage 3 and the storage time at stage 1 are estimated as 2.5 and 2.7 indicating that on average the output has approximately similar sensitivity to these factors. For these factors the probability of being statistically significant is 99 percent or more. However, although these factors have approximately similar average and range of rankings indicating that they are of comparable importance to each other, they are less important than the storage time at stage 3. The storage temperature at stage 1, lag period at stage 3, and lag period at stage 1 have mean ranks estimated as 4, 6, and 6.4, respectively. These factors can be considered as a group of moderate importance factors. The lag period at stage 2, the generation times at stages 1 and 2, the storage temperature at stage 2, and the maximum density have mean ranks between 9.2 and 9.7 with probability of being statistically significant varies between 5 to 13 percent indicating that the output shows the same sensitivity to this group of factors. The output has the lowest sensitivity to the storage time at stage 2 with mean rank of 12.1. This factor had statistically significant effect in only one of the uncertainty realizations.

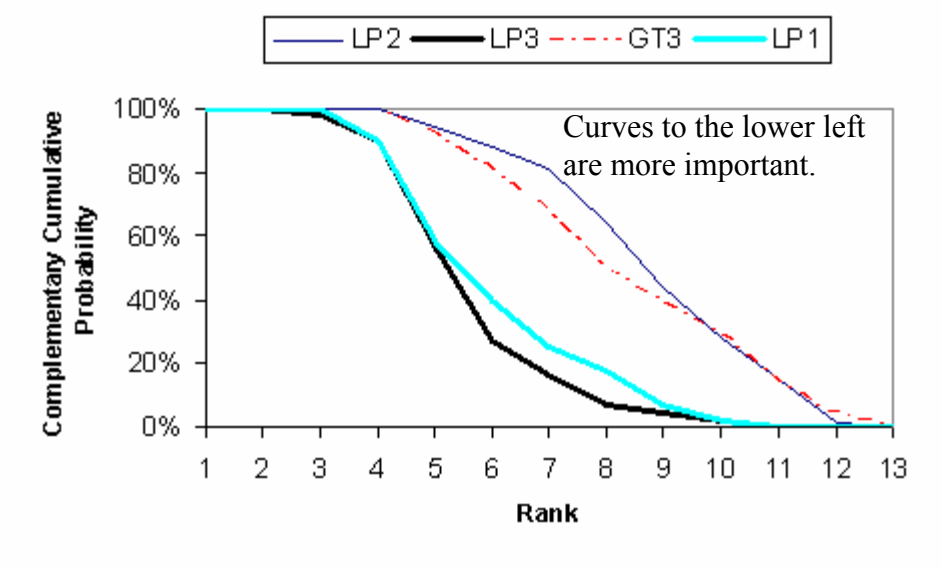



Figure 5-13. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected Factors: Lag Period at Stages 1, 2, and 3 ($LP_1$, $LP_2$, and $LP_3$); and Generation Time at Stage 3 ($GT_3$).

In order to visualize the results of the sensitivity analysis, the complementary cumulative distribution function (CCDF) of the rank is given for each factor based upon the 100 uncertainty realizations in Figures 5-12 to 5-14. Figure 5-12 displays the CCDFs for four factors that have the highest average ranks among all of the factors included in the analysis. These factors are storage time at stage 3, storage temperature at stage 3, storage time at stage 1, and storage temperature at stage 1. The CCDF for the storage time at stage 3 indicates that for 20 percent of the simulations, the rank was worse than one, which implies that the rank was equal to one for 80percent of the simulations. Furthermore, the storage temperature at stage 3 was ranked four or higher for 100 percent of the simulations. In contrast, storage time at stage 1 was ranked first for 10 percent of the simulations and was ranked fifth or higher for 98 percent of the simulations. The frequencies of being the most important factor for storage temperature at stages 1 and 3 are 5 and 7 percent, respectively. Thus, although the storage time at stage 3 has the highest frequency of a rank of one, there is some ambiguity regarding which of the other three factors is the second most important.

When comparing the CCDFs of Figure 5-12, it is apparent that the storage time at stage 3 tends to have a higher rank than the other factors. Furthermore, because the probability that the storage time at stage 3 has a rank of two or higher is 97 percent, the identification of the storage temperature as one of the most important factors is robust to uncertainty. In contrast, the storage

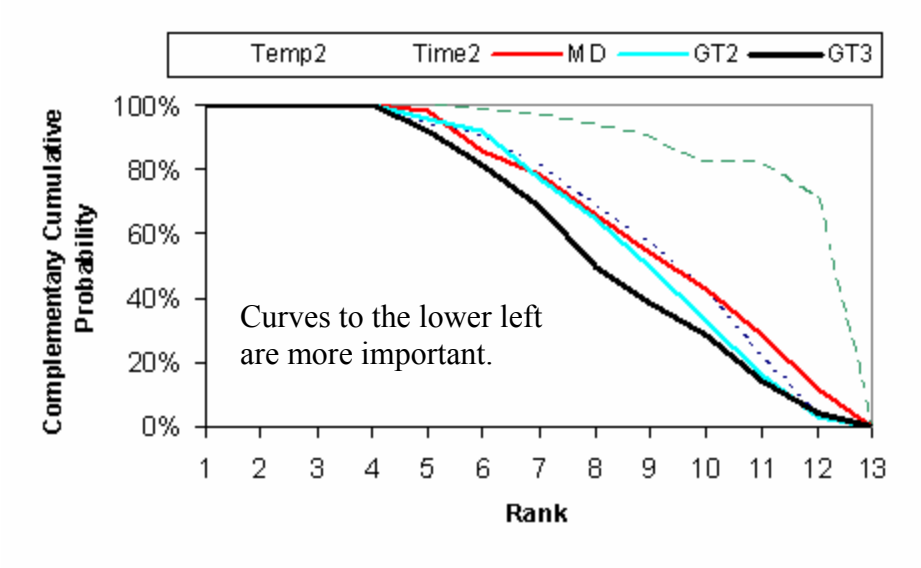



Figure 5-14. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected Factors: Storage Temperature and Time at Stage 2 ($Temp_2$ and $Time_2$); Maximum Density (MD); Generation Time at Stages 2 and 3 ($GT_2$ and $GT_3$).

time at stage 1 and the storage temperature at stages 1 and 3 have 32, 8, and 60 percent probability, respectively, of having a rank higher than two. Thus, these factors are typically less important than the storage time at stage 3.

Figure 5-13 displays the CCDFs for four factors that have the highest probability of a middle range of average ranks between six and nine among all of the factors included in the analysis. These factors are lag period at stages 1 to 3 ($LP_1$, $LP_2$, and $LP_3$) and generation time at stage 3 ($GT_3$). The CCDF for these factors indicate that for 100 percent of the simulations, the ranks for these factors were less than two. Although the mean ranks for these factors vary between six and nine, the probability that the ranks are worse than nine varies from 4 to 42 percent among the four selected factors. Furthermore, these four factors have ranks ranging from as high as 3 to as low as 13 in some uncertainty simulations. Thus, it is apparent that the identification of the rank of these factors is not robust to uncertainty. Therefore, there is ambiguity regarding the rank of each factor as a function of uncertainty in the model factors.

The least important group of factors is depicted in Figure 5-14. These factors include storage temperature and storage time at stage 2 ($Temp_2$ and $Time_2$), maximum density (MD), generation time at stage 2 ($GT_2$), and generation time at stage 3 ($GT_3$). These factors have a probability ranging from 83 to 100 percent of having a rank worse than five, and their average

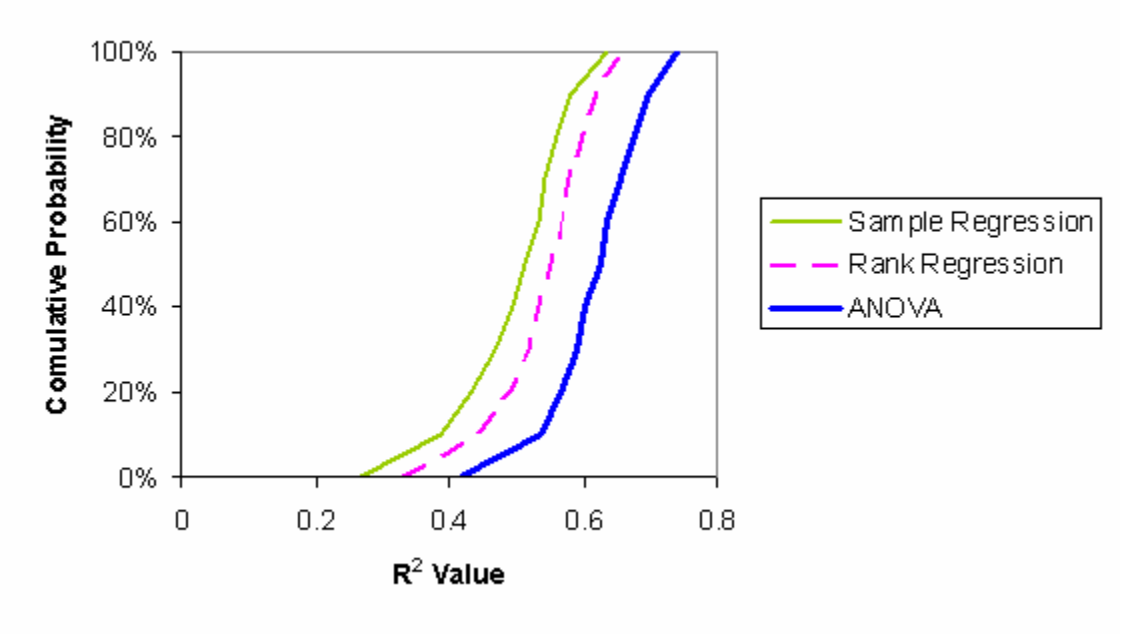



Figure 5-15. Comparison of the $R^2$ Distributions in Two-Dimensional Simulation of the Growth Estimation Part Based Upon Sample Regression, Rank Regression, and ANOVA.

ranks range from 10 to 13. MD, $GT_2$, and $Temp_2$ have similar CCDF distributions. The similarity of these distributions implies that these three factors are of comparable importance. There is ambiguity regarding the rank of each factor as a function of uncertainty in the model factors, although with high probability the ranks are worse than five in different uncertainty realizations. $Time_2$ can be identified as the least sensitive factor based on the CCDF distribution. $Time_2$ has a rank worse than 12 with probability of 72 percent in 100 uncertainty realizations. Moreover, this factor is almost statistically insignificant and just in one uncertainty iteration it was identified as a significant factor. Furthermore, even taking into account uncertainty, these five factors are clearly less important than the most important input, storage time at stage 3.

Figure 5-15 depicts the cumulative probability function (CDF) for the 100 $R^2$ values obtained in the two-dimensional simulation based upon ANOVA. The distributions of the $R^2$ values for the rank regression and standardized linear regression methods are also depicted in this figure in order to compare to that of ANOVA. Results of the rank regression and standardized linear regression are presented in Chapter 6. The $R^2$ value for ANOVA varied between 0.42 and 0.75 with an average of 0.63. This average is better that those obtained using standardized sample linear regression and rank regression, which had average $R^2$ values of 0.50 and 0.55, respectively. ANOVA does not impose any linearity assumption unlike linear

regression analysis. In this case, ANOVA includes only the main effect of the factors. These main effects account for average of 63 percent of the output variation. The larger $R^2$ values for ANOVA compared to the sample regression method implies that ANOVA is better able to respond to nonlinearities in the model. The larger $R^2$ values for ANOVA compared to the rank regression method implies that ANOVA is better able to respond to lack of monotonicity with respect to at least portions of the input domain. In particular, for some combinations of input values, growth is zero even though there is some variation in the inputs, or, alternatively, growth reaches a maximum and does not increase further even if some inputs increase. The average $R^2$ value for the results from ANOVA imply that approximately one third of the variance of the output is not explained by the selected factors and effects. Thus, opportunities may exist to increase the $R^2$ value by including additional interaction effects, such as third order effects. However, in this case, because the coefficient of determination implies that most of the variation in the output is accounted for, no additional refinement was made.

#### 5.4.1.3 One-Dimensional Simulation of Variability and Uncertainty

This section presents the results of ANOVA applied to a one-dimensional probabilistic simulation in which variability and uncertainty are co-mingled, based upon the case study scenario described in Section 3.3.3. The factor levels used in this analysis are the same as those given in Table 5-7.

The results of the analysis are given in Table 5-19. Rankings are presented for statistically significant factors with Pr>F less than 0.05. In addition to the main effect of each factor, the interaction effects between the storage time and the storage temperature at stages 1 to 3 are considered. F values in Table 5-19 indicate that there are no statistically significant effects for factors such as the storage time, the storage temperature, and the generation time at stage 2. Moreover, F values indicate that there is no statistically significant interaction between the storage time and the storage temperature at stage 2. Comparison of the F values for the interaction terms indicates that the interaction between the storage time and the storage temperature at stage 3 has higher importance that the interaction between these two factors at stage 1.

Comparing the magnitude of F values for the statistically significant factors indicates that the storage time at stage 3, the storage time at stage 1, the storage temperature at stage 3, and the storage temperature at stage 1 are the four most sensitive factors. The relative difference of F

Table 5-19. The Analysis of Variance Results for the Growth Estimation Part Based Upon One-Dimensional Co-mingled Variability and Uncertainty Simulation ($R^2 = 0.78$)

| Variable | F Value | Pr > F | Significant | Rank |
|---|---|---|---|---|
| Storage Temperature, Stage 1 ($Temp_1$) | 550 | <0.0001 | Yes | 4 |
| Storage Temperature, Stage 2 ($Temp_2$) | 0.1 | 0.99 | No | --- |
| Storage Temperature, Stage 3 ($Temp_3$) | 1,035 | <0.0001 | Yes | 3 |
| Storage Time, Stage 1 ($Time_1$) | 1,130 | <0.0001 | Yes | 2 |
| Storage Time, Stage 2 ($Time_2$) | 0.3 | 0.6 | No | --- |
| Storage Time, Stage 3 ($Time_3$) | 5,350 | <0.0001 | Yes | 1 |
| Maximum Density | 34 | <0.0001 | Yes | 9 |
| Lag Period, Stage 1 | 261 | <0.0001 | Yes | 5 |
| Lag Period, Stage 2 | 4 | 0.01 | Yes | 10 |
| Lag Period, Stage 3 | 200 | <0.0001 | Yes | 6 |
| Generation Time, Stage 1 | 52 | <0.0001 | Yes | 7 |
| Generation Time, Stage 2 | 1.0 | 0.03 | No | --- |
| Generation Time, Stage 3 | 45 | <0.0001 | Yes | 8 |
| $Temp_1$ * $Time_1$ | 1,190 | <0.0001 | Yes | |
| $Temp_2$* $Time_2$ | 0.04 | 0.96 | No | |
| $Temp_3$ * $Time_3$ | 2,270 | <0.0001 | Yes | |

values for these factors indicates that the ranking of the top factor is robust. The F values for the storage time at stages 3 and 1 differ by a ratio of 4.7 indicating that the rank of the storage time at stage 3 is robust. In contrast, comparison of the F values for the storage time at stage 1 and the storage temperature at stage 3 indicates that the relative ranking of these two factors are not robust. The F values for these factors differ by a ratio of only 1.1. The robustness of these rankings was further evaluated in Section 5.4.1.2, when the factors were ranked for different uncertainty realizations.

For factors rather than the four most sensitive ones, small F values indicate that these factors may be unimportant. For example, F values for the generation time at stage 3, the maximum density, and the lag period at stage 2 which are ranked eight to ten, differ by ratios 12 to 138 compared to the F value of the fourth important factor.

The coefficient of determination, $R^2$, is 0.78 indicating that a substantial proportion of the variation in the output is captured by including the main effects and interactions between selected factors. Although including additional effects, such as additional two way interactions or higher order interactions, may increase the coefficient of determination, the coefficient of determination based upon this analysis is sufficiently high to confirm that most of the variation

Table 5-20. Evaluation of ANOVA Contrasts Regarding the Interactions Between the Storage Temperature and the Storage Time at Stage 1

| Contrast | Estimate | F Value | Pr>F | Significant |
|---|---|---|---|---|
| T <7.5-11>$^{o}$C, Time <0-24> and <24-48> hour | 0.005 | 112.21 | <0.0001 | Yes |
| T <7.5-11>$^{o}$C, Time <24-48> and <48-72> hour | 0.026 | 1283.0 | <0.0001 | Yes |
| T <7.5-11>$^{o}$C, Time <48-72> and <72-96> hour | 0.049 | 1721.0 | <0.0001 | Yes |
| T <7.5-11>$^{o}$C, Time <72-96> and <96-120> hour | 0.069 | 1278.0 | <0.0001 | Yes |
| T <7.5-11>$^{o}$C, Time <96-120> and <120-144> hour | 0.074 | 610.0 | <0.0001 | Yes |
| T <7.5-11>$^{o}$C, Time <120-144> and <144-168> hour | 0.103 | 455.0 | <0.0001 | Yes |
| T <7.5-11>$^{o}$C, Time <144-168> and <168-192> hour | 0.042 | 30.8 | <0.0001 | Yes |
| T <7.5-11>$^{o}$C, Time <168-192> and <192-216> hour | 0.031 | 6.8 | 0.008 | Yes |
| T <7.5-11>$^{o}$C, Time <192-216> and <216-240> hour | 0.108 | 38.8 | <0.0001 | Yes |
| T <7.5-11>$^{o}$C, Time <216-240> and <240-264> hour | ------ | 0.16 | 0.8 | No |
| T <11-14.5>$^{o}$C, Time <0-24> and <24-48> hour | 0.116 | 4062.0 | <0.0001 | Yes |
| T <11-14.5>$^{o}$C, Time <24-48> and <48-72> hour | 0.211 | 5290.0 | <0.0001 | Yes |
| T <11-14.5>$^{o}$C, Time <48-72> and <72-96> hour | 0.218 | 2168.0 | <0.0001 | Yes |
| T <11-14.5>$^{o}$C, Time <72-96> and <96-120> hour | 0.119 | 241.0 | <0.0001 | Yes |
| T <11-14.5>$^{o}$C, Time <96-120> and <120-144> hour | 0.087 | 48.4 | <0.0001 | Yes |
| T <11-14.5>$^{o}$C, Time <120-144> and <144-168> hour | ------ | 0.9 | 0.6 | No |
| T <18-21.5>$^{o}$C, Time <0-24> and <24-48> hour | 0.55 | 2630.0 | <0.0001 | Yes |
| T <18-21.5>$^{o}$C, Time <24-48> and <48-72> hour | ----- | 0.1 | 0.4 | No |
| T <21.5-25>$^{o}$C, Time <0-24> and <24-48> hour | 0.503 | 6270 | <0.0001 | Yes |
| T <21.5-25>$^{o}$C, Time <24-48> and <48-72> hour | ----- | 2.4 | 0.09 | No |

in the output is captured. Therefore, in this case, no additional refinement was made to the analysis.

In addition to the inferences obtained by application of ANOVA regarding the sensitivity of the output to individual factors, additional information regarding sensitivity is achieved by using contrasts. Contrasts were prepared only for the mixed analysis, since this analysis takes into account the widest range of possible values of each factor in the context of a single simulation. The selection of factors to include in the contrasts was based upon the sensitivity results of the individual factors. The rankings in Table 5-19 indicate that the storage time and the storage temperature at stages 3 and 1 are statistically significant. Hence, a set of contrasts was prepared to evaluate the response of the model to the interaction between these factors. The results of these contrasts are summarized in Tables 5-20 and 5-21.

Contrasts in Tables 5-20 and 5-21 compare the mean response in the growth estimation part, considering the interaction between the storage temperature and the storage time at stages 1

Table 5-21. Evaluation of ANOVA Contrasts Regarding the Interactions Between the Storage Temperature and the Storage Time at Stage 3

| Contrast | Estimate | F Value | Pr>F | Significant |
|---|---|---|---|---|
| T <7.5-11>$^{o}$C, Time <0-24> and <24-48> hour | 0.003 | 30 | <0.0001 | Yes |
| T <7.5-11>$^{o}$C, Time <24-48> and <48-72> hour | 0.013 | 227 | <0.0001 | Yes |
| T <7.5-11>$^{o}$C, Time <48-72> and <72-96> hour | 0.032 | 381 | <0.0001 | Yes |
| T <7.5-11>$^{o}$C, Time <72-96> and <96-120> hour | 0.055 | 351 | <0.0001 | Yes |
| T <7.5-11>$^{o}$C, Time <96-120>$^{r}$ and <120-144> hour | 0.081 | 196 | <0.0001 | Yes |
| T <7.5-11>$^{o}$C, Time <120-144> and <144-168> hour | 0.0.4 | 18.5 | 0.0007 | Yes |
| T <7.5-11>$^{o}$C, Time <144-168>$^{r}$ and <168-192> hour | ------ | 0.61 | 0.5 | No |
| T <11-14.5>$^{o}$C, Time <0-24> and <24-48> hour | 0.104 | 5,122 | <0.0001 | Yes |
| T <11-14.5>$^{o}$C, Time <24-48> and <48-72> hour | 0.198 | 5,748 | <0.0001 | Yes |
| T <11-14.5>$^{o}$C, Time <48-72> and <72-96> hour | 0.186 | 1,524 | <0.0001 | Yes |
| T <11-14.5>$^{o}$C, Time <72-96> and <96-120> hour | 0.084 | 101 | <0.0001 | Yes |
| T <11-14.5>$^{o}$C, Time <96-120> and <120-144> hour | ------ | 3.4 | 0.08 | No |
| T <14.5-18>$^{o}$C, Time <0-24> and <24-48> hour | 0.55 | 3,374 | <0.0001 | Yes |
| T <14.5-18>$^{o}$C, Time <24-48> and <48-72> hour | 0.256 | 1,239 | <0.0001 | Yes |
| T <14.5-18>$^{o}$C, Time <48-72> and <72-96> hour | 0.129 | 88 | 0.0006 | Yes |
| T <14.5-18>$^{o}$C, Time <72-96> and <96-120> hour | ----- | 0.7 | 0.8 | No |
| T <18-21.5>$^{o}$C, Time <0-24> and <24-48> hour | 0.421 | 18,169 | <0.0001 | Yes |
| T <18-21.5>$^{o}$C, Time <24-48> and <48-72> hour | 0.068 | 143 | <0.0001 | Yes |
| T <18-21.5>$^{o}$C, Time <48-72> and <72-96> hour | 0.035 | 9 | 0.03 | Yes |
| T <18-21.5>$^{o}$C, Time <72-96> and <96-120> hour | ----- | 0.7 | 0.8 | No |
| T <21.5-25>$^{o}$C, Time <0-24> and <24-48> hour | 0.433 | 4,933 | <0.0001 | Yes |
| T <21.5-25>$^{o}$C, Time <24-48> and <48-72> hour | ----- | 2.7 | 0.5 | No |

and 3. The 'Estimate' columns in Tables 5-20 and 5-21 present the estimate of the difference between the mean responses for the condition mentioned in the contrast. If the estimate is not significant as indicated by Pr>F greater 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 5-20 indicate that when the storage temperature in stage 1 is at the first level (e.g., between 7.5$^{o}$C and 11$^{o}$C) the storage time does matter in the growth of the *E. coli* organisms in the ground beef servings until the tenth day. After the tenth day there is no significant difference between the estimated growth in coming days indicating that the tenth day is a saturation time for the growth of the *E. coli* organisms. When the storage temperature at stage 1 is at the second level (e.g., between 11$^{o}$C and 14.5$^{o}$C) the storage time matters only until the fifth day, indicating that with an increase in the storage temperature the growth of the *E. coli* organisms reaches the saturation point in only five days. When the storage temperature at stage 1

Table 5-22. Summary of the ANOVA Results for Growth Estimation Part Based on Variability Only, Variability for Different Uncertainty Realizations, and Co-mingled Variability and Uncertainty Analyses

| Variable | Ranks | | |
|---|---|---|---|
| | **Analysis 1(1)** | **Analysis 2(2)** | **Analysis 3(3)** |
| Storage Temperature, Stage 1 | 3 | 4.0 | 4 |
| Storage Temperature, Stage 2 | 10 | 9.6 | --- |
| Storage Temperature, Stage 3 | 2 | 2.5 | 3 |
| Storage Time, Stage 1 | 4 | 2.7 | 2 |
| Storage Time, Stage 2 | --- | 12.1 | --- |
| Storage Time, Stage 3 | 1 | 1.3 | 1 |
| Maximum Density | 9 | 9.7 | 9 |
| Lag Period, Stage 1 | 5 | 6.4 | 5 |
| Lag Period, Stage 2 | 11 | 9.2 | 10 |
| Lag Period, Stage 3 | 6 | 6 | 6 |
| Generation Time, Stage 1 | 8 | 9.3 | 7 |
| Generation Time, Stage 2 | --- | 9.3 | --- |
| Generation Time, Stage 3 | 7 | 8.8 | 8 |

(1) Ranks based on the variability only analysis.
(2) Mean ranks based on the variability for different uncertainty realizations analysis.
(3) Ranks based on the one-dimensional co-mingled variability and uncertainty analysis.

increases to the third, forth and fifth levels, the saturation point is reached in four, three, and two days, respectively. This pattern implies that with increase in the storage temperature, the saturation time for the growth of the *E. coli* organisms is reached in shorter time.

Table 5-21 demonstrates the same pattern for the interaction between the storage temperature and the storage time at stage 3. When the storage temperature in stage 3 is at its first level (e.g., between 7.5$^{o}$C and 11$^{o}$C), the saturation time for the growth is reached in seven days. An increase in the storage temperature at this stage causes the saturation time for the growth of the *E. coli O157:H7* organisms to happen faster. For example, when the storage temperature is at its second level, the saturation happens in only five days. The saturation in the growth of *E. coli* organisms occurs after 4, 4, and 2 days when the storage temperature at stage 3 is at its third, fourth and fifth levels, respectively.

#### 5.4.1.4 Summary and Comparison of the Results of ANOVA in the Growth Estimation Part

In Sections 5.4.1.1 to 5.4.1.3 ANOVA was applied to three datasets considering variability only, variability for different uncertainty realizations, and co-mingled variability and

uncertainty in factors. In this section rankings based on these analyses are summarized and compared. Table 5-22 gives the ranks for each factor based on analyses in Sections 5.31 to 5.3.3. Table 5-22 indicates that the storage time at stage 3 is identified as the most important factor based upon all three simulations. The storage temperature at stage 3 was identified as the second most important factor based upon the variability only simulation and as the third most important factor based upon co-mingling of variability and uncertainty. However, for both of these simulations, the F value for this factor was not substantially different than the F-value for the next ranked factor. Furthermore, the group of top four factors was the same for both simulation methods. Moreover, in the second analysis taking into account the uncertainty in factors, the storage time and the storage temperature at stages 3 and 1 are identified as the top four important factors. Hence, all three simulations have agreement on the top four sensitive factors. The variability only and the co-mingled variability and uncertainty analyses identified the storage time and the generation time at stage 2 as statistically insignificant factors. These factors have average ranks of 12.1 and 9.3 with the second analysis with frequency of being statistically significant of only 1 and 9 percent, respectively.

All three approaches presented in Sections 5.4.1.1 to 5.4.1.3 yielded similar rankings with respect to the most important factor, a group of three factors of secondary importance, a group of five factors with minor importance, and a group of three factors as unimportant.

### 5.4.2 Uncertainties in Estimates of F Values

The objective of this section is to evaluate the uncertainty associated with point estimates of F values, such as those produced in earlier sections of this chapter. Because the F value is estimated based upon a random sample of values for inputs to the model, the F value is itself a random variable. Thus, a key question is regarding how much the F values of two inputs must differ in order to infer that the two inputs have substantially different importance with regard to sensitivity. Because the procedure for estimating uncertainty in F values is computationally intensive, it was applied to only one case study for a selected part of the *E. coli* model. Specifically, the variability only simulation of the growth estimation part of the preparation module was chosen.

The method of bootstrap simulation was 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). There are several variants of bootstrap

Table 5-23. Summary of the ANOVA Results for 200 Bootstrap Simulations for F value Statistics

| Variable | Mean F Value | 95% Probability Range | SD/Mean | Frequency (Percent) | Var. Rank (1) | Mean Rank | Range of Rank |
|---|---|---|---|---|---|---|---|
| Temp1 | 481 | (318,606) | 0.16 | 100 | 3 | 4.0 | 3-4 |
| Temp2 | 1.8 | (0.0,9.6) | 1.45 | 17 | 10 | 10.9 | 7-13 |
| Temp3 | 1010 | (810,1180) | 0.09 | 100 | 2 | 1.0 | 1-2 |
| Time1 | 657 | (557,780) | 0.09 | 100 | 4 | 2.9 | 2-4 |
| Time2 | 0.4 | (0.0,2.6) | 1.58 | 1 | NS (2) | 12.3 | 9-13 |
| Time3 | 781 | (714,915) | 0.06 | 100 | 1 | 2.1 | 1-3 |
| MD | 8.6 | (1.3,26) | 0.71 | 79 | 9 | 8.1 | 7-13 |
| LP1 | 50 | (35,64) | 0.14 | 100 | 5 | 5.9 | 5-6 |
| LP2 | 1.5 | (0.1,5.0) | 0.76 | 15 | 11 | 10.6 | 9-13 |
| LP3 | 60 | (47,73) | 0.11 | 100 | 6 | 5.1 | 5-6 |
| GT1 | 16 | (9.3,24) | 0.26 | 100 | 8 | 7.2 | 7-8 |
| GT2 | 1.7 | (0.1,4.8) | 0.74 | 17 | NS (2) | 9.9 | 8-12 |
| GT3 | 19 | (12,25) | 0.21 | 100 | 7 | 6.8 | 6-8 |

(1) Ranks based on variability only analysis from Table 5-17
(2) Identified as not statistically significant in the variability only analysis

simulation. In this case, an empirical bootstrap method was used. For the variability only simulation, 65,000 random values were generated for each factor based upon specified probability distribution models. In the empirical bootstrap approach, an alternative randomized version of the original Monte Carlo simulation is obtained by sampling with replacement from the original 65,000 random values. This procedure is computationally faster than generating a new random sample of 65,000 from the original specified probability distribution models. In order to estimate confidence intervals, it is typically desirable to simulate thousands of bootstrap samples. However, because of the computational requirements for each bootstrap sample, it was feasible to simulate only 200 bootstrap samples. Therefore, 200 bootstrap samples, each based upon 65,000 samples with replacement, were generated for all model factors. ANOVA was applied to each of the bootstrap samples to produce a distribution of 200 F values for each factor.

The bootstrap simulation results are summarized in Table 5-23. These results indicate that there is a substantial range of uncertainty associated with the estimates of the 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 a range of approximately plus or minus 20 percent of the mean value. The storage time at

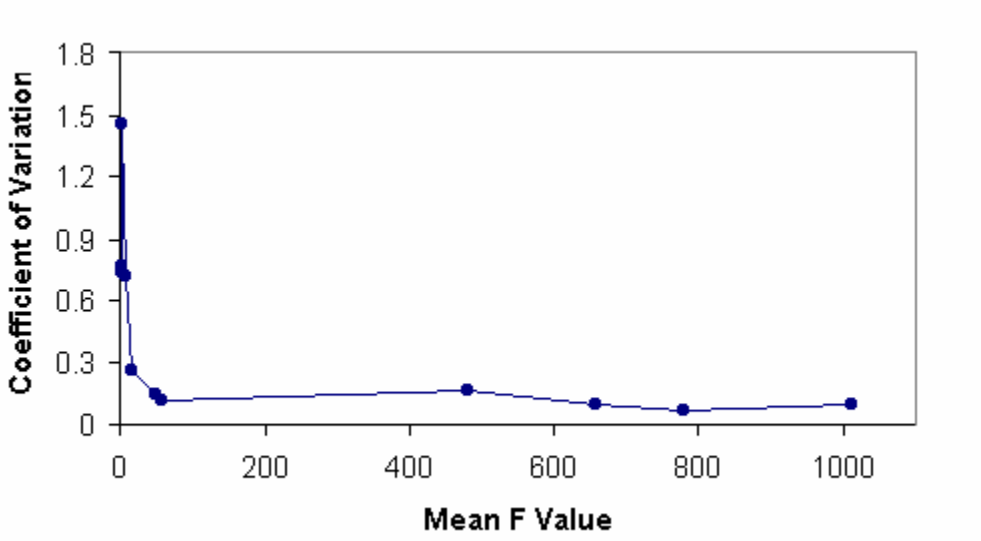



Figure 5-16. Coefficient of Variation Versus the Mean for Bootstrap F Values.

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. As will be shown later, the sampling distributions of these two F values are approximately independent of each other. Therefore, the overlap in their confidence intervals indicates a possibility that the rank order between these two inputs can reverse, even though on average the F value for the storage temperature is larger than that for the storage time by a factor of 1.3. However, the storage temperature at stage 3 has a statistically significantly larger F value than the factor with the third highest average rank, which is the storage time at stage 1. The probability ranges for F values of these two factors do not overlap. Therefore, although there is some ambiguity regarding which of two inputs may be the most important, it is clear that the storage temperature at stage 3 is more important than the storage time at stage 1.

In fact, it is possible to clearly distinguish several groups of inputs. The first group includes storage temperature and time at stage 3, with mean F values of 781 to 1,010 and 95 percent probability range enclosing values from 714 to 1,180. The second group includes storage temperature and time at stage 1, with mean F values of 481 to 657, and intervals enclosing values from 318 to 780. The third group includes the lag periods at stages 1 and 3, with mean F values of 50 to 60 with intervals enclosing values from 35 to 74. Thus, the third group is clearly less important than the second group. The fourth group includes mean F values

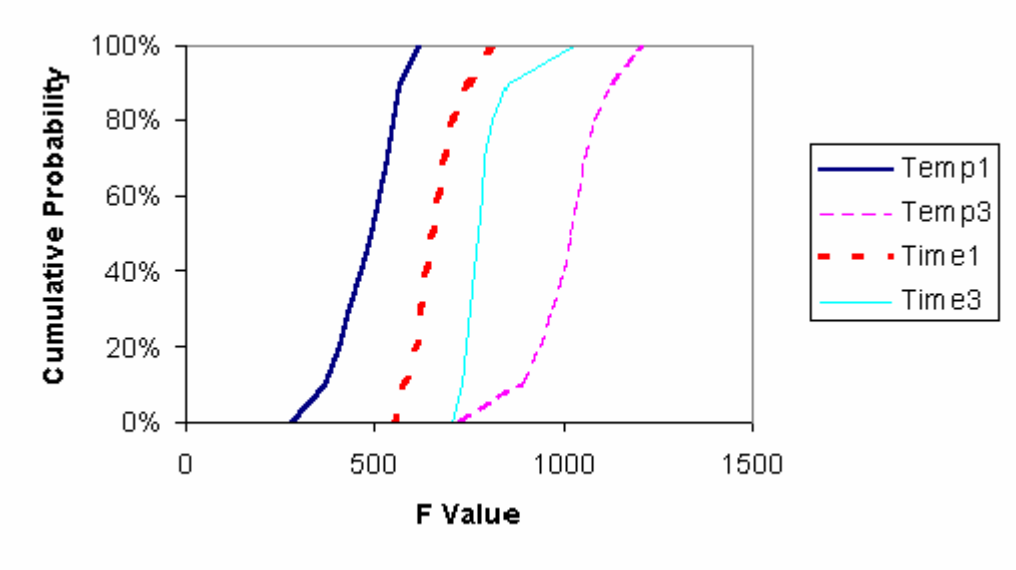



Figure 5-17. Cumulative Distribution Function of the Bootstrap F Values for Selected Factors: Storage Temperature and Storage Time at Stages 1 and 3 (Temp1, Time1, Temp3, and Time3).

from approximately 9 to 19, with intervals including values as low as 1.3 and as high as 26. The fifth and final group includes mean F values of less than 2.0 with intervals typically from 0.0 to as much as 10. In this latter group most of the bootstrap samples produced statistically insignificant F values. With the exception of some overlap in the intervals between factors in the first and second groups, the intervals among the groups typically do not overlap.

The results from the bootstrap simulation are comparable in many ways to the results obtained from point-estimates of F values from the original variability only simulation. In particular, both analyses produced similar rank ordering for groups of factors. Although the numerical values of the ranks from the variability only simulation often do not agree with the average ranks from the bootstrap simulation, the differences can be attributed to random sampling error and the resulting ambiguity in ranks within groups of factors. For example, the bootstrap simulation results imply that there can be reversals in the rank order of the top two inputs. Thus, although the top ranked input from the variability only analysis was for the factor that had a mean rank of 2.1 in the bootstrap simulation, the apparent difference in the rank is not statistically significant. The variability only analysis correctly identified the storage time and temperature at stage 1 as less important than the top two factors and as more important than all other factors. The difference in rank order between the storage time and temperature at stage 1 when comparing the variability only analysis with the bootstrap simulation is attributable to

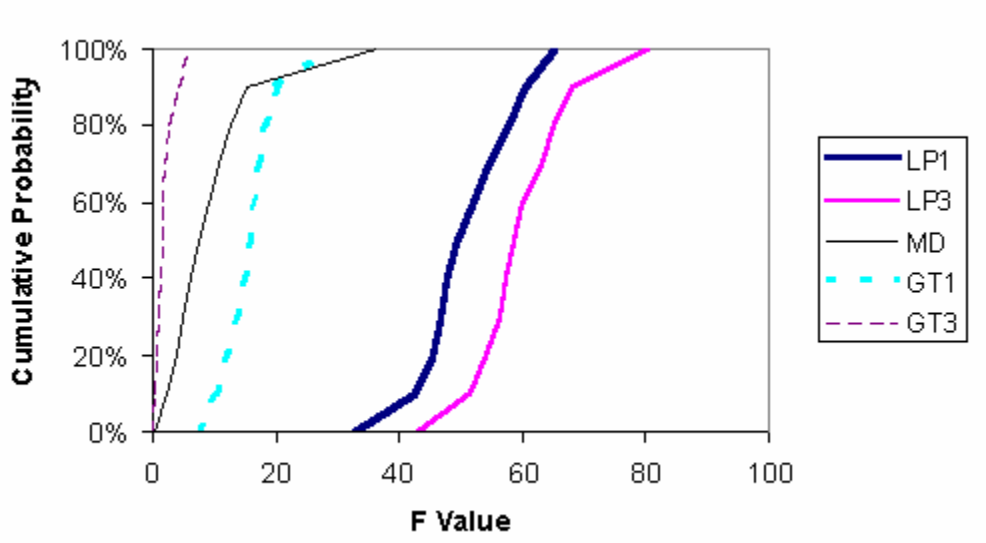



Figure 5-18. Cumulative Distribution Function of the Bootstrap F Values for Selected Factors: Lag Period and Generation Time at Stages 1 and 3 (LP1, GT1, LP3, and GT3).

random sampling error, since the probability ranges for the F values of these two factors overlap considerably. The third group of factors has similar mean F values and similar probability ranges; therefore, either of the two factors in this group could be ranked fifth or sixth. The factors identified as ranked seventh through ninth in the variability only analysis correspond to the fourth group identified from the bootstrap simulation. The factors ranked tenth or lower, including statistically insignificant factors, correspond to the fifth group identified based upon the bootstrap simulation.

In previous sections, relative differences in F values were used to make a judgment as to whether two F values were substantially different from each other. 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 substantially different, the range of the sampling distribution of the F values must be considered. A compact method for visualizing the range of uncertainty in the F value is to plot the coefficient of variation, which is the standard deviation divided by the mean, versus the mean value. Figure 5-16 shows the coefficient of variation versus the mean value based upon the bootstrap results of Table 5-23. The coefficient of variation is approximately 0.15 or less for F values greater than approximately 50. For F values smaller than approximately 20, the coefficient of variation ranges from approximately 0.3 to 1.6. These results suggest that the

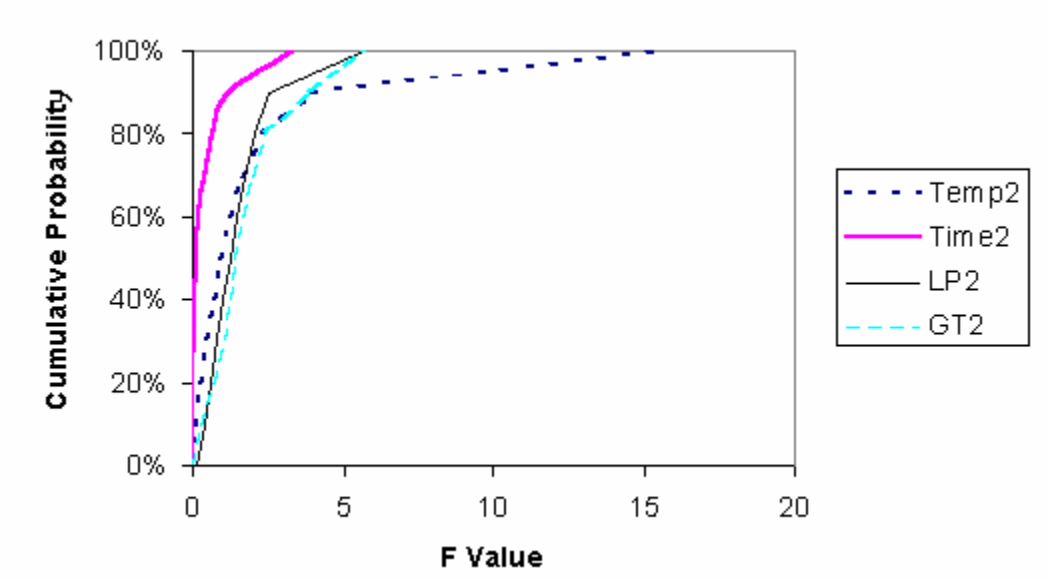



Figure 5-19. Cumulative Distribution Function of the Bootstrap F Values for Selected Factors: Storage temperature, Storage Time, Lap Period, and Generation Time at Stage 2 (Temp2, Time2, LP2, and GT2); and Maximum Density (MD).

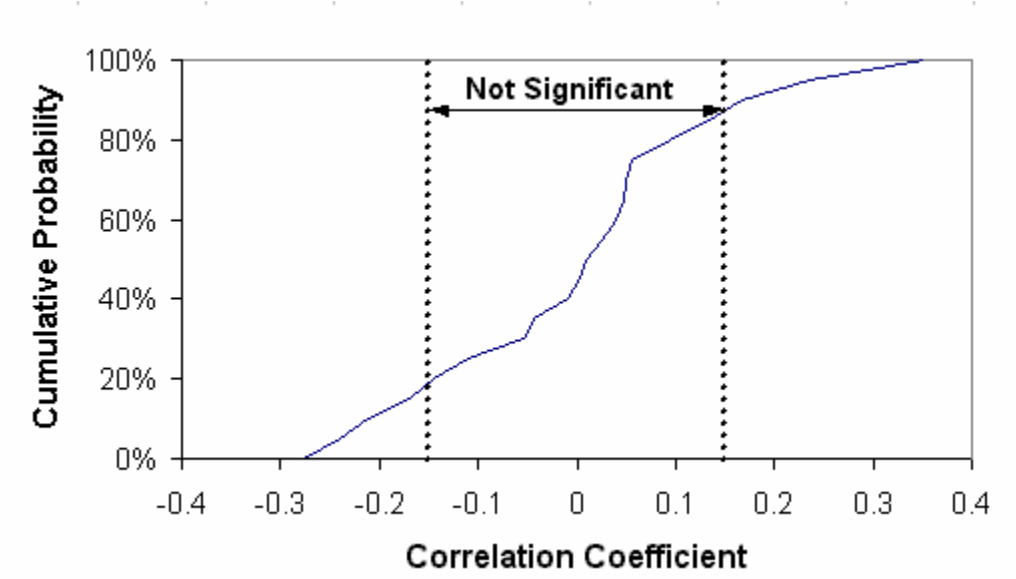



Figure 5-20. Cumulative Density Function for Correlation Coefficients Between F Values.

Table 5-24. Correlation Coefficients Between Bootstrap F values for Inputs to the Growth Estimation Part

| | Temp1 | Temp2 | Temp3 | Time1 | Time2 | Time3 | MD | LP1 | LP2 | LP3 | GT1 | GT2 | GT3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Temp1** | 1.000 | 0.144 | -0.172 | 0.000 | 0.010 | 0.022 | 0.162 | -0.277 | -0.036 | 0.049 | -0.053 | 0.056 | -0.051 |
| **Temp2** | | 1.000 | -0.004 | 0.073 | -0.050 | -0.069 | 0.041 | -0.025 | -0.206 | 0.035 | -0.154 | 0.248 | -0.150 |
| **Temp3** | | | 1.000 | -0.188 | -0.079 | 0.142 | 0.272 | -0.197 | 0.050 | -0.273 | 0.025 | 0.182 | 0.028 |
| **Time1** | | | | 1.000 | -0.026 | -0.002 | -0.223 | 0.230 | -0.275 | 0.032 | -0.227 | 0.127 | -0.237 |
| **Time2** | | | | | 1.000 | 0.009 | -0.075 | 0.146 | -0.162 | 0.233 | 0.045 | -0.044 | 0.041 |
| **Time3** | | | | | | 1.000 | 0.130 | 0.005 | 0.140 | 0.040 | 0.050 | 0.073 | 0.047 |
| **MD** | | | | | | | 1.000 | -0.132 | 0.352 | -0.128 | -0.239 | 0.054 | -0.259 |
| **LP1** | | | | | | | | 1.000 | -0.165 | 0.107 | 0.010 | -0.199 | 0.012 |
| **LP2** | | | | | | | | | 1.000 | -0.240 | 0.047 | 0.182 | 0.042 |
| **LP3** | | | | | | | | | | 1.000 | 0.028 | -0.232 | 0.026 |
| **GT1** | | | | | | | | | | | 1.000 | -0.120 | -0.125 |
| **GT2** | | | | | | | | | | | | 1.000 | -0.126 |
| **GT3** | | | | | | | | | | | | | 1.000 |

Table 5-25. P Values for Estimated Correlation Coefficients Between Bootstrap F values for Inputs to the Growth Estimation Part

| | Temp1 | Temp2 | Temp3 | Time1 | Time2 | Time3 | MD | LP1 | LP2 | LP3 | GT1 | GT2 | GT3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Temp1** | | 0.042 | 0.015 | 0.996 | 0.885 | 0.754 | 0.022 | <0.0001 | 0.613 | 0.490 | 0.458 | 0.427 | 0.460 |
| **Temp2** | | | 0.959 | 0.307 | 0.486 | 0.334 | 0.565 | 0.726 | 0.003 | 0.618 | 0.029 | 0.000 | 0.031 |
| **Temp3** | | | | 0.008 | 0.266 | 0.048 | <0.0001 | 0.005 | 0.483 | <0.0001 | 0.724 | 0.010 | 0.702 |
| **Time1** | | | | | 0.712 | 0.982 | 0.002 | 0.001 | <0.0001 | 0.657 | 0.001 | 0.073 | 0.001 |
| **Time2** | | | | | | 0.897 | 0.289 | 0.039 | 0.022 | 0.001 | 0.526 | 0.533 | 0.536 |
| **Time3** | | | | | | | 0.067 | 0.941 | 0.048 | 0.572 | 0.486 | 0.305 | 0.491 |
| **MD** | | | | | | | | 0.063 | <0.0001 | 0.072 | 0.001 | 0.450 | 0.001 |
| **LP1** | | | | | | | | | 0.020 | 0.131 | 0.890 | 0.005 | 0.891 |
| **LP2** | | | | | | | | | | 0.001 | 0.512 | 0.010 | 0.517 |
| **LP3** | | | | | | | | | | | 0.697 | 0.001 | 0.699 |
| **GT1** | | | | | | | | | | | | 0.091 | 0.089 |
| **GT2** | | | | | | | | | | | | | 0.090 |
| **GT3** | | | | | | | | | | | | | |

coefficient of variation may be relatively constant for F values that are statistically significant and substantially large. Since a 95 percent probability range might typically be enclosed by plus or minus two standard deviations for a symmetric sampling distribution, one might infer in this case that statistically significant F values that differ by 30 percent or more would typically be associated with clear differences in the rank order of the corresponding factors. These results are specific to this model and the simulation sample size of 65,000 used for the Monte Carlo simulation and should not be used to make quantitative judgments of the significance in differences between F values obtained based upon other sample sizes or models.

The sampling distributions for the F values are illustrated in Figures 5-17 through 5-19 as cumulative distribution functions (CDFs). Figure 5-17 shows the two groups of the most important factors. The distribution for the F values of storage temperature at stage 3 has the largest F values but overlaps substantially with the distribution of F values for the storage time at stage 3. Although there is a small amount of overlap of the distribution of F values for storage time at stage 1 with that of storage time at stage 1, for the most part these two distributions are substantially different. The distribution of F values for storage temperature at stage 1 does not overlap with any of the other three and has smaller F values; therefore, this distribution is said to be stochastically dominated by the others.

Figure 5-19 displays the third and fourth groups of factors. The CDFs within each group overlap substantially, but the CDFs for factors in one group do not overlap with the CDFs of factors in the other group. Thus, the distribution of F values in the third group stochastically dominated the distribution of F values in the fourth group. Figure 5-19 displays the CDFs for the fifth group, which includes a high proportion of statistically insignificant F values. These four factors are of approximately comparable unimportance.

When comparing distributions of F values that have substantial overlap, it is important to know whether the distributions are statistically independent. In order to evaluate the independence of the sampling distributions of the F values, a correlation matrix was estimated based upon the 200 bootstrap replications of the F values for all 13 factors, as shown in Table 5-24. The P-values corresponding to each correlation coefficient are given in Table 5-25. Correlation coefficients that have a magnitude of less than 0.142 are not statistically significant.

Correlation coefficients between -0.142 and 0.142 are deemed to be not statistically significantly different from zero and are indicative of statistical independence. The range of

Table 5-26. The Analysis of Variance Results for the Cooking Effect Part

| Variable | F Value | Pr > F | Significant | Rank |
|---|---|---|---|---|
| Precooking Treatment (PCT) | 8,400 | <0.0001 | Yes | 2 |
| Cooking Place ($C_p$) | 4 | 0.07 | No | --- |
| Cooking Temperature ($C_T$) | 10,900 | <0.0001 | Yes | 1 |
| PCT * $C_p$ | 1 | 0.4 | No | |
| PCT * $C_T$ | 380 | <0.0001 | Yes | |
| $C_p$ * $C_T$ | 2 | 0.2 | No | |

correlation coefficients in Table 5-23 is depicted as a CDF in Figure 5-20. Approximately 65 percent of the estimated correlation coefficients are not statistically significant. Approximately 10 percent of the correlations are larger than 0.25. The largest magnitudes of the estimated correlation coefficients are approximately 0.3. Although such values are statistically significant, they are nonetheless weak correlations. Thus, a reasonable approximation is that the sampling distributions of the bootstrap F values are independent or that any correlation between them is very weak.

The main methodological findings of this analysis are as follows: (1) the sampling distributions of F values as quantified based upon 200 bootstrap simulations each with a sample size of 65,000 have a coefficient of variation of approximately 0.15 for large average F values; (2) differences in F values of approximately 30 percent or more imply a clear discrimination in rank order; and (3) the sampling distributions of F values are approximately independent of each other. The main case study-specific findings are that it was possible to separate the 13 inputs into five groups in which several factors within a group were of comparable importance. These groups were similar to those obtained based upon point estimation of F values for a single Monte Carlo analysis. Although the development of bootstrap sampling distributions of F values is computationally intensive, this is a useful method for gaining insight into the statistical significance of differences between F values. The results obtained here are specific to the simulation sample size of 65,000. For smaller simulation sample sizes, the distribution of F values is expected to be larger; therefore, the relative difference between F values associated with statistically significant differences in ranks would be larger than for the case study presented here.

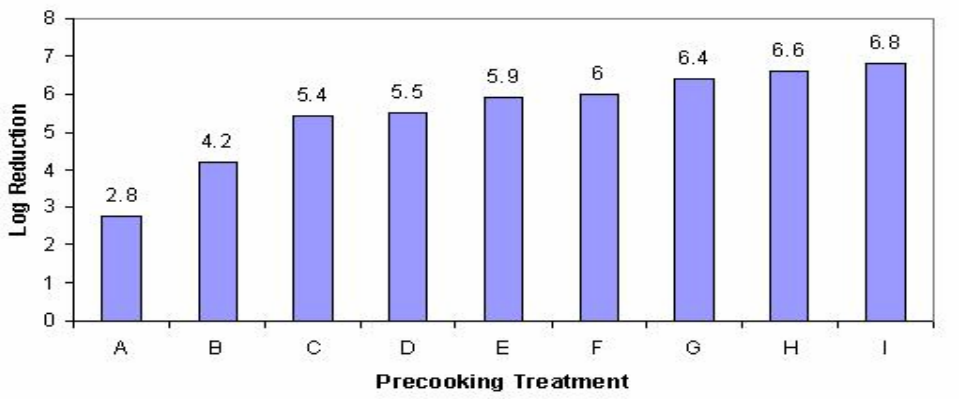



Figure 5-21. Mean Log Reduction in the Number of *E. coli* Organisms Due to the Cooking Effect for Different Precooking Treatments.

### 5.4.3 Analysis of Variance for Variability in the Cooking Effect Part

As explained in Section 3.4.3.2, factors for the cooking effect part include cooking temperature, precooking treatment, and cooking place. Distributions for these factors are summarized in Table 3-13. The output in the cooking effect part is the mean log reduction in the number of *E. coli* organisms. In Section 5.1.3, the definition of levels for each factor is explained and in Table 5-6 the assigned levels are summarized. For the cooking effect part there is a one-dimensional variability simulation with 65,000 iterations. Table 5-26 summarizes the result of application of ANOVA to the cooking effect part.

The factors in Table 5-26 are ranked based on the magnitude of F values. Rankings are presented for statistically significant factors with Pr>F less than 0.05. Rankings are presented considering the F values only for main effects. In addition to the main effect of each factor, the interaction effect between precooking treatment and cooking place, between precooking treatment and cooking temperature, and between cooking place and cooking temperature are also considered in the model. The F values in Table 5-26 indicate that cooking place is statistically insignificant. The interaction effect between the precooking treatments and cooking temperature is statistically significant; however, other interaction terms have no significant effects. The cooking temperature is the most sensitive factor. Hence, it is ranked first. The precooking treatment has a rank of two based on the magnitude of the F value. The difference between the F values of the cooking temperature and precooking treatment indicates that the rankings for these top two factors may not be robust, since their F values differ by a ratio of only 1.3.

In order to better understand the relationship between the mean response and levels of the precooking treatment, the mean response is estimated for each level of this factor in Figure 5-21. Different levels for the precooking treatment were defined in Table 3-8. The mean value of the log reduction in the number of *E. coli* organisms is highest for the precooking treatment *I* and is approximately 2.4 times greater than the value estimated based upon the precooking treatment *A*. Moreover, precooking treatments *G*, *H*, and *I* have approximately the same mean responses.

### 5.4.4 Analysis of Variance for Variability in the Serving Contamination Part

As explained in section 3.4.3.1, factors for the serving contamination part include the ground beef consumption type, serving size, eating location, consumer age, and grinder contamination. Distributions for these factors are summarized in Table 3-12. The output in this part is the mean serving contamination. In Section 5.1.3, the definition of levels for each factor is explained and in Table 5-6 the assigned levels are summarized. For this part there is a one-dimensional variability simulation with 65,000 iterations, as explained in Section 3.3.3. The case scenario in the serving contamination part focuses on the high and low prevalence seasons, separately. Tables 5-27 and 5-28 summarize the result of application of ANOVA to the serving contamination part in high and low prevalence seasons, respectively.

Factors in Tables 5-27 and 5-28 are ranked based on the magnitude of F values. Rankings are presented for the main effects of statistically significant factors with Pr>F less than 0.05. In addition to the main effect for each factor, the interactions between these factors are also considered in the model. F values in Table 5-27 indicate that the consumer age does not have a statistically significant effect in the high prevalence season, although its interactions with other factors such as the serving size, the ground beef consumption type, and the eating location are significant. F values in Table 5-28 indicate that the consumer age and the eating location are statistically insignificant in the low prevalence season. The serving size is the most sensitive factor in both high and low prevalence seasons. The grinder contamination, and the ground beef consumption type are ranked second and third, respectively, in high and low prevalence seasons. The difference between the F values indicates that the rank of the first sensitive factor is robust. The F value for the serving size is approximately 11 times greater than the F value for the grinder contamination in both high and low prevalence seasons.

F values for the interaction terms indicate that these interactions have statistically significant effects in both the high and low prevalence seasons. The interaction between the

Table 5-27. The Analysis of Variance Results for the Serving Contamination in High Prevalence Season

| Variable | F Value | Pr > F | Significant | Rank |
|---|---|---|---|---|
| Ground Beef Consumption Type (GBT) | 122 | <0.0001 | Yes | 3 |
| Eating Location (Loc) | 6 | 0.02 | Yes | 4 |
| Consumer Age ($C_a$) | 2 | 0.2 | No | --- |
| Serving Size (S) | 9,400 | <0.0001 | Yes | 1 |
| Grinder Contamination ($G_{con}$) | 820 | <0.0001 | Yes | 2 |
| $C_a$ * S | 16 | <0.0001 | Yes | |
| Loc * S | 700 | <0.0001 | Yes | |
| GBT * S | 35 | <0.0001 | Yes | |
| GBT * $C_a$ | 30 | <0.0001 | Yes | |
| GBT * $G_{con}$ | 211 | <0.0001 | Yes | |
| Loc * $C_a$ | 62 | <0.0001 | Yes | |
| GBT * Loc | 250 | <0.0001 | Yes | |

Table 5-28. The Analysis of Variance Results for the Serving Contamination in Low Prevalence Season

| Variable | F Value | Pr > F | Significant | Rank |
|---|---|---|---|---|
| Ground Beef Consumption Type (GBT) | 105 | <0.0001 | Yes | 3 |
| Eating Location (Loc) | 2 | 0.2 | No | --- |
| Consumer Age ($C_a$) | 2 | 0.2 | No | --- |
| Serving Size (S) | 9,200 | <0.0001 | Yes | 1 |
| Grinder Contamination ($G_{con}$) | 925 | <0.0001 | Yes | 2 |
| $C_a$ * S | 42 | <0.0001 | Yes | |
| Loc * S | 825 | <0.0001 | Yes | |
| GBT * S | 60 | <0.0001 | Yes | |
| GBT * $C_a$ | 35 | <0.0001 | Yes | |
| GBT * $G_{con}$ | 170 | <0.0001 | Yes | |
| Loc * $C_a$ | 80 | <0.0001 | Yes | |
| GBT * Loc | 220 | <0.0001 | Yes | |

eating location and the serving size has the highest F value in both high and low prevalence seasons indicating that this interaction is most important. F value for the interaction between the eating location and the serving size differs from the second important interaction term based on the magnitude of the F value by ratios of 2.8 and 3.8 for high and low prevalence seasons, respectively.

In addition to the inferences obtained by application of ANOVA regarding the sensitivity of the output to individual factors, additional information regarding sensitivity is achieved by

Table 5-29. Contrasts for Checking the Interaction Effects in the Serving Contamination Part

| Contrast | Estimate | F Value | Pr>F |
|---|---|---|---|
| Hamburger vs Eating Location (Home or Away) in Winter | 0.0001 | 131 | <0.0001 |
| Hamburger vs Eating Location (Home or Away) in Summer | 0.0001 | 114 | <0.0001 |
| Meatball vs Eating Location (Home or Away) in Winter | -0.0001 | 16 | <0.0001 |
| Meatball vs Eating Location (Home or Away) in Summer | -0.0001 | 24 | <0.0001 |
| Hamburger vs Age ({5-24} &{25-64}) in Winter | 0.0001 | 137 | <0.0001 |
| Hamburger vs Age ({5-24} &{25-64}) in Summer | 0.00007 | 60 | <0.0001 |
| Hamburger vs Age ({25-64} &{64+}) in Winter | -0.0001 | 33 | <0.0001 |
| Hamburger vs Age ({25-64} &{64+}) in Summer | -0.0001 | 34 | <0.0001 |
| Eating Location (home) vs Serving Size (g) ({120-150}&{150-80}) in Winter | 0.0006 | 116 | <0.0001 |
| Eating Location (home) vs Serving Size (g) ({120-150}&{150-180}) in Summer | 0.0006 | 87 | <0.0001 |
| Eating Location (away) vs Serving Size (g) ({120-150}&{150-180}) in Winter | 0.0007 | 148 | <0.0001 |
| Eating Location (away) vs Serving Size (g) ({120-150}&{150-180}) in Summer | 0.0007 | 124 | <0.0001 |
| Eating Location (home) vs Serving Size (g) ({150-180)&{>180}) in Winter | 0.0025 | 2096 | <0.0001 |
| Eating Location (home) vs Serving Size (g) ({150-180)&{>180}) in Summer | 0.0023 | 1597 | <0.0001 |
| Eating Location (away) vs Serving Size (g) ({150-180)&{>180}) in Winter | 0.0038 | 3860 | <0.0001 |
| Eating Location (away) vs Serving Size (g) ({150-180)&{>180}) in Summer | 0.0033 | 2897 | <0.0001 |

using contrasts. Contrasts are useful in order to find thresholds in the model response to different factors, or in understanding the response of the model to interactions between factors.

Table 5-29 summarizes the results of contrasts in the serving contamination part considering the interactions between factors involving in the simulation. The 'Estimate' column in Table 5-29 presents the estimate of the difference between the mean responses for the condition given in the 'Contrast' column. If the estimate is not significant, which means that the Pr>F is greater than 0.05, there is not enough statistical support indicating that the estimated value for the contrast is different from zero. All the estimates in Table 5-29 are statistically significant. According to the contrasts in Table 5-26, higher contamination for hamburger patties is expected away from home

in comparison with those servings made at home during both the high and low prevalence seasons. The positive estimate for this contrast in Table 5-29 indicates that the contamination in hamburger patties away from home is higher compared to hamburger patties at home. This result can be justified based on the idea that servings away from home are made from grinders coming from combo bins, while some of the home servings are produced from grinders coming from trim boxes. Grinders coming from trim boxes have lower contamination than those grinders from combo bins (FSIS, 2001). Moreover, according to the contrasts in Table 5-29, meatballs at home are more contaminated than those meatball servings away from home during high and low prevalence seasons. For meatballs, unlike the hamburger patties, the interaction between the grinder beef consumption type and the eating location is dominated by the serving size, which is bigger at home than away from home (Table 3-7).

Contrasts in Table 5-29 also indicate that hamburger patties consumed by people between 25 and 64 years old are more contaminated than those hamburger patties consumed by people between 5 and 24. Moreover, hamburger patties eaten by consumers above 65 have lower contamination than those patties eaten by people between 25 and 64. The larger serving size for people between 25 and 64 years old is consistent with this conclusion. Based on contrast results in Table 5-29, with increase in the serving size for any type of ground beef consumed at home or away from home during high and low prevalence seasons, an increase in the mean serving contamination is expected.

## 5.5 Evaluation of ANOVA as a Sensitivity Analysis Method Based on Applications to the *E. coli* Model

In this chapter ANOVA was applied to specific modules and parts of the *E. coli* model in order to identify the most important factors influencing the response of selected outputs. In some cases, ANOVA was applied to the same part based upon three different types of probabilistic analysis, including simulation of variability only, variability and uncertainty in two dimensions, and variability and uncertainty co-mingled in one dimension.

The slaughter module and the growth estimation part had a two-dimensional variability and uncertainty characteristic that made it possible to implement three different types of probabilistic analysis. The results from all three approaches were typically comparable in terms of the rank ordering of inputs or the identification of groups of inputs of similar importance. This result is likely to be specific to the case studies evaluated here.

ANOVA is able to deal with categorical factors. Continuous factors had to be converted into discrete ranges, referred to as levels. The assignment of levels to a factor is a matter of judgment. A trade-off in selecting levels is regarding the number of levels and the number of data points included within a given level. Each method of level definition has its advantages and disadvantages. In this chapter three methods were used in order to define levels for each factor: (1) equal intervals; (2) equal percentiles; and (3) visual inspection of the CDF for each factor.

Defining levels with equal intervals helps in identifying the possible threshold in the model response. For example, in Section 5.4.1.3 equal levels were considered for the storage time and the storage temperature at stages 1 and 3 in order to estimate the saturation point for the growth of *E. coli* organisms in ground beef servings. The saturation points were identified using contrast in ANOVA. With an increase in the number of levels for these factors and the use of similar intervals, the saturation point can be estimated with more accuracy. Considering equal percentiles for definition of levels guarantees an equal number of data points in each level, thereby leading to a balanced experiment with estimable contrasts. Using the CDF of each factor in level definition facilitates the evaluation of the model response in the lower or upper tail of the factor distribution. For example, in Section 5.3.3 this method was used to define the levels for the Trim/Vacuum/Washing efficiency. Using contrasts from ANOVA, the sensitivity of combo bin contamination to high and low decontamination efficiencies corresponding to upper and lower tails of this factor were evaluated for different levels of the chilling process. The results from the contrasts can be implemented in decision-making regarding practical approaches to decrease the amount of contamination in the slaughter plants. For example, if there is insufficient control regarding the storage time and the storage temperature during the chilling process, more attention should be paid to the decontamination step. With high efficiency during the decontamination process, using Trim/Vacuum/wash, it is possible to decrease approximately 2.6 logs in the contamination of the combo bins as the final product of the slaughter plants.

The application of ANOVA to identify the importance of the interactions among factors was demonstrated in this chapter. For factorial experiments with contributions of more than one factor in the model in addition to the simple effect of each factor, the interaction between factors should be considered in ANOVA. These interaction terms can be compared to each other based on the magnitude of the estimated F values.

Three case studies are provided in the growth estimation part in which the coefficients of determination, $R^2$, are presented for ANOVA. The $R^2$ values in these cases indicate that the models incorporating the main effects of the factors and interaction effects between selected factors captured a substantial high amount of variation in the output. For the two-dimensional probabilistic approach, a comparison was made between the cumulative probability distribution of $R^2$ values obtained in ANOVA with those of standardized linear regression and rank regression analysis. This comparison indicated that on average ANOVA captured a higher proportion of variation in the output comparing to the other two methods. This finding implies that classification of the range of each input to the factor levels performed in ANOVA did not deteriorate the amount of variability in the output that could be captured.

The uncertainty in point estimates of F values should be taken into account when making comparisons of the F values of two or more factors. For a Monte Carlo simulation sample size of 65,000 with a particular model, the range of uncertainty in statistically significant F values that were substantially large was found to be approximately plus or minus 30 percent or less. This implies that differences in F values of 30 percent or more for a simulation sample size of 65,000 are associated with clear differences in rank order between factors. In situations where the F values are similar, factors can be categorized into groups of similar importance. It is also possible to discriminate between groups of factors such that there are clear differences between groups. Therefore, ANOVA is a reasonable method for characterizing the sensitivity of model inputs and it can deal with nonlinearities, thresholds, and categorical inputs.

# 6 REGRESSION ANALYSIS FOR THE *E. COLI O157:H7* MODEL

The objective of this chapter is to evaluate regression analysis and related techniques as methods for sensitivity analysis based upon application to the *E. coli* food safety risk assessment model. The specific methods evaluated here include sample (Pearson) correlation coefficients, rank (Spearman) correlation coefficients, linear sample-based regression, and rank regression. The details of these methods are discussed in Chapter 2. Although the use of linear regression is the main focus of this chapter, the correlation and rank regression methods are included in selective case studies to enable comparisons among these methods. For example, sample and rank correlations are commonly used by practitioners because these methods are often included in commercial software packages, such as Crystal Ball, that are used in many risk assessment studies.

This chapter contains three parts, presenting results of the application of regression analysis to the production, slaughter, and preparation modules. In the production module, the feedlot prevalence, within feedlot prevalence, breeding herd prevalence, and within breeding herd prevalence parts are analyzed separately in Section 6.1. In the slaughter module three analyses are performed based upon different methods for quantifying variability and uncertainty as discussed in Section 6.2. In the preparation module, the growth estimation, cooking effect, and the serving contamination parts are analyzed separately as described in Section 6.3. The growth estimation part is selected for analyses with the rank regression and correlation coefficients methods. Results from these two methods are compared with the results obtained from standardized linear regression analysis. Based on the comparison, key similarities and differences between these methods are identified. In Section 6-4, regression analysis is evaluated as a method for sensitivity analysis and the limitations, advantages, disadvantages and key criteria for application of this method are summarized. Moreover, an evaluation is given for the rank regression and correlation coefficient methods for sensitivity analysis.

As explained in Section 2.2.1, results from standardized linear regression are sensitive to the linear assumption regarding the relationship between the output and the inputs to the model. Hence, for this method of analysis the coefficient of determination, $R^2$, is provided for each analysis. The $R^2$ value indicates the percent of variation in the model response that is explained by the inputs considered in the linear model. A high $R^2$ value implies that the linearity assumption for the functional relation between the output and inputs is substantially valid, while

low $R^2$ values indicate deviation from the underlying linearity assumption. Results of the sensitivity analysis in cases with high $R^2$ are considered to be reasonably valid.

Regression analysis can be applied to models that include both quantitative (e.g., continuous) and qualitative (e.g., categorical) inputs. However, as discussed in Section 2.2.1, instead of a single coefficient for each qualitative input, a set of coefficients is estimated for corresponding indicator variables. These coefficients cannot be compared with those of the quantitative inputs. Therefore, a recommended approach is to use the F value associated with the qualitative and quantitative inputs, rather than the regression coefficients, as a basis for rank ordering the importance among the inputs. Furthermore, the importance of the qualitative inputs can be assessed using a graphical approach. The use of F values and graphics to support inferences regarding the importance of qualitative inputs is demonstrated in this chapter for the production module and the serving contamination part of the preparation module.

## 6.1 Regression Analysis in the Production Module

In the production module, regression analysis is applied to four parts, including the feedlot prevalence, within feedlot prevalence, breeding herd prevalence, and within breeding herd prevalence. The results of the analyses for these four parts are presented in Sections 6.1.1 to 6.1.4, respectively.

### 6.1.1 Uncertainty in the Feedlot Prevalence Part

As described in Section 3.2.1, for feedlot prevalence estimation, the inputs include the apparent prevalence and the herd sensitivity as quantitative inputs, and the study as a qualitative one. The output of interest in the feedlot prevalence part is the median feedlot prevalence. Distributions for the inputs in this part are given in Table 3-9. In order to address qualitative variables in regression analysis, quantitative indicators for different classes of the qualitative inputs are employed (Neter *et al.*, 1996). The indicator variables are frequently also called dummy variables. Using the CLASS statement in SAS© facilitates definition of the dummy variables for inputs (SAS, 1996).

The case scenario in Section 3.3.1 includes a one-dimensional uncertainty simulation with 65,000 iterations. Equation 2-6 is used to normalize the generated data in the feedlot prevalence part for the herd sensitivity and the apparent prevalence as quantitative inputs. Table 6-1 summarizes the result of the regression analysis in the feedlot prevalence part.

Table 6-1. Regression Analysis Results for Sensitivity Analysis of Uncertainty in the Feedlot Prevalence Part of the Production Module ($R^2 = 0.82$)

| Variable | Coefficient | 95% CI [(a)] | F Value | Pr>F | Rank |
|---|---|---|---|---|---|
| Study | --- | --- | 22,800 | <0.0001 | --- |
| Apparent Prevalence | 0.06 | (0.02, 0.10) | 8 | 0.006 | 2 |
| Herd Sensitivity | -0.25 | -0.25[(b)] | 9,430 | <0.0001 | 1 |

(a) CI = Confidence Interval for the coefficient
(b) The interval for this coefficient is so tight that it appears as -0.25 to -0.25 when it is rounded to two decimal places.

The rankings in Table 6-1 are based on the magnitude of the estimated standardized regression coefficients for quantitative inputs. Rankings are presented for statistically significant inputs with Pr>F less than 0.05. The effect of different study levels is to shift the fitted linear model up or down based on the change caused by the specific study level.

Figure 6-1 illustrates the effect of the study on the fitted linear model. The linear relation between the output and the herd sensitivity is plotted for each study level. The median feedlot prevalence is more sensitive to the choice of study than to the herd sensitivity. For example, for a herd sensitivity of 0.5, the median feedlot prevalence varies from approximately 60 to 95 percent depending upon the choice of study, or a range of approximately 35 percentage points.

In contrast, for a given choice of study, such as Hancock (1998), the median feedlot prevalence varies between approximately 50 and 65 percent, or a range of approximately 15 percentage points. Thus, the typical range of variation in the median feedlot prevalence is much wider with respect to the choice of study than it is with respect to the value for the herd sensitivity.

The F values in Table 6-1 indicate that all inputs are statistically significant. Rankings based on the coefficient estimates indicate that the median feedlot prevalence is most sensitive to the herd sensitivity, because this input has a coefficient with a larger magnitude than those of other inputs. Hence, the herd sensitivity is ranked first. The apparent prevalence is ranked second.

The 95 percent confidence intervals are estimated for quantitative inputs in order to evaluate how clear the rankings are. Comparison of the 95 percent confidence intervals for the herd sensitivity and the apparent prevalence indicates that the ranking for the herd sensitivity as the first ranked input is unambiguous, because the confidence intervals for this input do not overlap with those of the second ranked input.

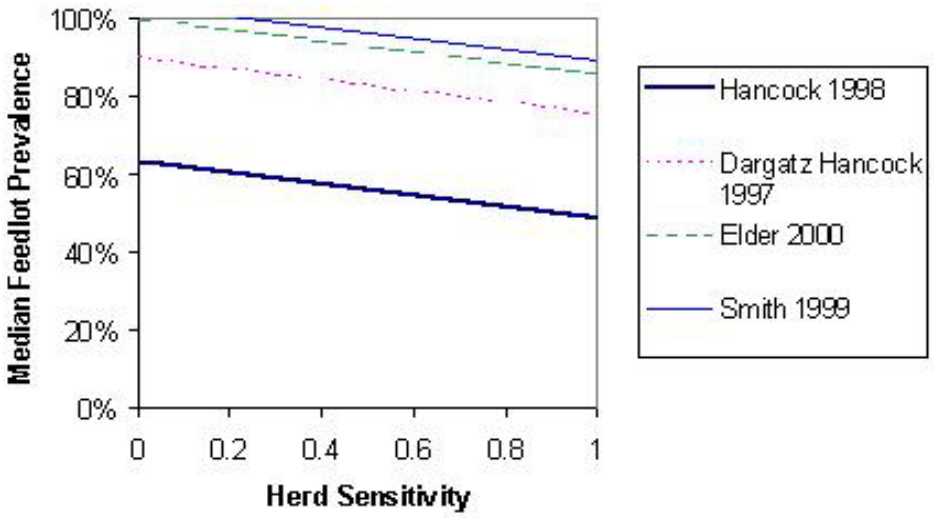



Figure 6-1. Regression Lines for Different Study Levels in the Feedlot Prevalence Part.

The use of F values instead of coefficients to gain insight into key inputs would lead to different rankings. If instead of the coefficient estimates, the magnitude of the F values is used as a criterion for ranking the inputs, the study would be ranked as the most important input, while the herd sensitivity and the apparent prevalence would be placed as the second and third important inputs, respectively.

The use of F values as a method for ranking sensitive inputs that include both qualitative and quantitative values is compared to the graphical results shown in Figure 6-1. The graphical results imply that the choice of study has a more substantial impact on the median feedlot prevalence than does the value for the herd sensitivity. The study has an F value that is approximately a ratio of 2.4 larger than that for the herd sensitivity. Thus, the comparison of F values indicates that the effect of the study is stronger than that of herd sensitivity. Therefore, the use of F values as a means for comparison of the importance of qualitative inputs versus quantitative inputs appears to have intuitive appeal.

The coefficient of determination, $R^2$, for the linear regression model fitted to the dataset is 0.82 indicating that the linear assumption for the functional relation between the output and inputs is substantially valid.

Table 6-2. Regression Analysis Results for Sensitivity Analysis of Uncertainty in the Within Feedlot Prevalence Part of the Production Module (R2 = 0.90)

| Variable | Coefficient | 95% CI [a] | F Value | Pr>F | Rank |
|---|---|---|---|---|---|
| Study | --- | --- | 910 | <0.0001 | --- |
| Season | --- | --- | 300 | <0.0001 | --- |
| Apparent Within Feedlot Prevalence | 1.50 | (1.49, 1.51) | 12,700 | <0.0001 | 1 |
| Test Sensitivity | -0.22 | -0.22[b] | 8,700 | <0.0001 | 2 |

(a) CI = Confidence Interval for the coefficient.
(b) The interval for this coefficient is so tight that it appears as -0.22 to -0.22 when it is rounded to two decimal places.

### 6.1.2 Uncertainty in Within Feedlot Prevalence Part

Section 3.2.1 explains the within feedlot prevalence part. The key inputs for this part include the apparent within feedlots prevalence and the test sensitivity as quantitative inputs, and the study and the season as qualitative inputs. Table 3-9 summarizes the distributions for these inputs. The output of interest is the average within feedlot prevalence. The case scenario for this part is based upon a one-dimensional uncertainty simulation with 65,000 iterations as described in Section 3.3.1. Equation 2-6 is used to normalize the Monte Carlo simulation data for the quantitative inputs. Table 6-2 summarizes the result of the regression analysis in this part.

The rankings in Table 6-2 are based on the magnitude of the estimated regression coefficients for quantitative inputs. Rankings are presented for statistically significant inputs with Pr>F less than 0.05. F values in Table 6-2 indicate that all inputs are statistically significant. Rankings based on the coefficient estimates indicate that the average within feedlot prevalence is most sensitive to the apparent within feedlot prevalence. Hence, the apparent within feedlot prevalence is ranked first. The test sensitivity is ranked second. The study and the season are not considered in this ranking, because no coefficient is estimated for these inputs.

The 95 percent confidence intervals are estimated for quantitative inputs. There is no overlap for the estimated confidence intervals for the quantitative inputs. Therefore, the rankings are considered unambiguous.

If instead of the coefficient estimates, the magnitude of the F values is used as a criterion for ranking the inputs, the apparent within feedlots prevalence, the test sensitivity, the study and the season would be ranked first to fourth, respectively. Figure 6-2 illustrates the effect of the study on the fitted linear model. The linear relationship between the output in summer and the

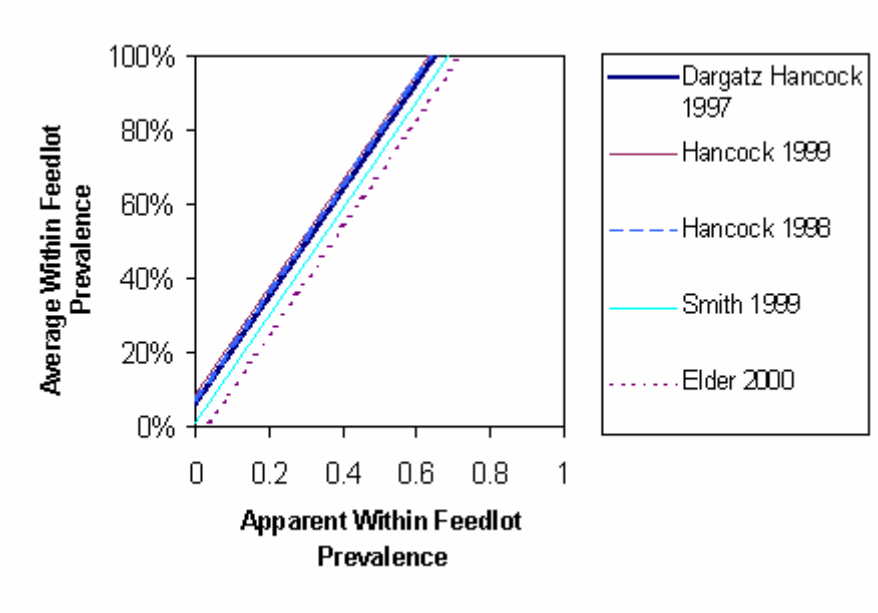



Figure 6-2. Regression Lines for Different Study Levels in the Within Feedlot Prevalence Part in Summer.

apparent within feedlot prevalence is plotted for each study level. The output is more sensitive to the apparent within feedlot prevalence than to the choice of the study level. For example, for an apparent within feedlot prevalence of 0.5, the output varies from approximately 68 to 80 percent depending upon the choice of study, or a range of approximately 12 percentage points. In contrast, for a given choice of study, such as Hancock (1999), the output varies between approximately 10 and 100 percent, or a range of approximately 90 percentage points. Thus, the typical range of variation in the average within feedlot prevalence is much wider with respect to value for the apparent within feedlot prevalence than it is with respect to the choice of study.

The $R^2$ for the linear regression model fitted to the dataset is 0.90, which is quite high. Thus, the linear assumption for the functional relationship between the output and inputs appears to be reasonable.

### 6.1.3 Uncertainty in the Breeding Herd Prevalence Part

As described in Section 3.2.1, for the breeding prevalence estimation, apparent prevalence and the herd sensitivity are quantitative inputs and the study is a qualitative input. The output is the median breeding herd prevalence. Distributions for the inputs are given in Table 3-9. The case scenario in Section 3.3.1 is based upon a one-dimensional uncertainty

Table 6-3. Regression Analysis Results for Sensitivity Analysis of Uncertainty in the Breeding Herd Prevalence Part of the Production Module ($R^2 = 0.90$)

| Variable | Coefficient | 95% CI [(a)] | F Value | Pr>F | Rank |
|---|---|---|---|---|---|
| Study | --- | --- | 15,600 | <0.0001 | --- |
| Apparent Prevalence | -0.04 | (-0.32,0.24) | 0.1 | 0.8 | --- |
| Herd Sensitivity | -0.20 | -0.20[(b)] | 3,600 | <0.0001 | 1 |

(a) CI = Confidence Interval for the coefficient
(b) The interval for this coefficient is so tight that it appears as -0.20 to -0.20 when it is rounded to two decimal places.

simulation with 65,000 iterations. Equation 2-6 is used to normalize the generated data in the breeding herd prevalence part for the herd sensitivity and the apparent prevalence as quantitative inputs. Table 6-3 summarizes the result of the regression analysis in the breeding herd prevalence part.

The rankings in Table 6-3 are based on the magnitude of the estimated regression coefficients for quantitative inputs. Rankings are presented for significant inputs with Pr>F less than 0.05. The F values in Table 6-3 indicate that all inputs except the apparent prevalence are statistically significant. Rankings based on the magnitude of the coefficient estimates indicate that the median breeding herd prevalence is most sensitive to the herd sensitivity.

If instead of the coefficient estimates, the magnitude of the F values is used as a criterion for ranking the inputs, the study would be ranked as the most important input, while the herd sensitivity would be placed as the second important inputs.

Figure 6-3 illustrates the effect of the study on the fitted linear model. The linear relation between the output and the herd sensitivity is plotted for each study level. The median breeding herd prevalence is more sensitive to the choice of study than to the herd sensitivity. For example, for a herd sensitivity of 0.5, the median breeding herd prevalence varies from approximately 40 to 100 percent depending upon the choice of study, or a range of approximately 60 percentage points. In contrast, for a given choice of study, such as Garber (1998), the median breeding herd prevalence varies between approximately 30 and 58 percent, or a range of approximately 28 percentage points. Thus, the typical range of variation in the median breeding herd prevalence is much wider with respect to the choice of study than it is with respect to the value for the herd sensitivity. Hence, the graphical results imply that the choice of study has a more substantial impact on the median breeding herd prevalence than does the value for the herd sensitivity. The comparison of the F values for these two inputs reveals that the

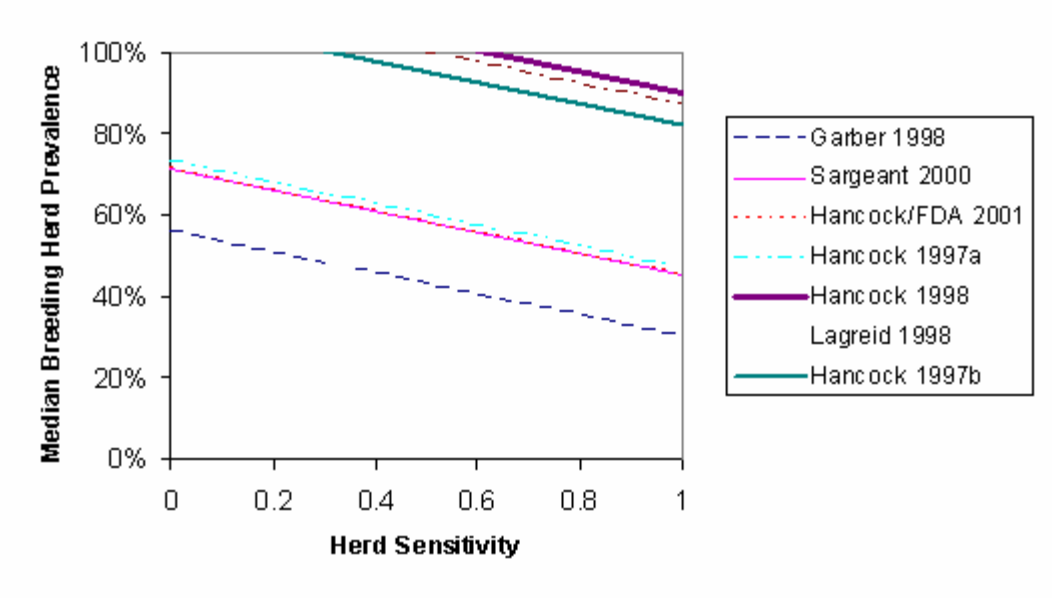



Figure 6-3. Regression Lines for Different Study Levels in the Breeding Herd Prevalence Part.

study has an F value that is approximately a ratio of four larger than that for the herd sensitivity. Thus, the comparison of F values indicates that the effect of the study is stronger than that of herd sensitivity. Therefore, the use of F values as a means for comparison of the importance of qualitative inputs versus quantitative inputs appears to have intuitive appeal.

The $R^2$ for the linear regression model fitted to the dataset is 0.90. This high value of $R^2$ implies that the linear assumption for the functional relation between the output and inputs is substantially valid.

### 6.1.4 Uncertainty in the Within Breeding Herd Prevalence Part

Section 3.2.1 explains the within breeding herd prevalence part. The apparent within breeding herd prevalence and test sensitivity are quantitative inputs, while study and season are qualitative inputs. Table 3-9 summarizes the distributions for these inputs. The output is the average within breeding herd prevalence. The case scenario for this part is based upon a one-dimensional uncertainty simulation with 65,000 iterations. Equation 2-6 is used to normalize the generated data in this part for quantitative inputs. The results of the regression analysis in this part are given in Table 6-4.

The rankings in Table 6-4 are based on the magnitude of the estimated regression coefficients for quantitative inputs. These rankings are only presented for statistically significant inputs with Pr>F less than 0.05. The F values indicate that there is not a statistically significant

Table 6-4. Regression Analysis Results for Sensitivity Analysis of Uncertainty in the Within Breeding Herd Prevalence Part of the Production Module ($R^2$ = 0.84)

| Variable | Coefficient | 95% CI [a] | F Value | Pr>F | Rank |
|---|---|---|---|---|---|
| Study | --- | --- | 2,500 | <0.0001 | --- |
| Season | --- | --- | 0.2 | 0.8 | --- |
| Apparent Within Breeding Herd Prevalence | 1.17 | (1.13,1.20) | 3,150 | <0.0001 | 1 |
| Test Sensitivity | -0.30 | (-0.32,-0.28) | 2,100 | <0.0001 | 2 |

(a) CI = Confidence Interval for the coefficient

effect for the season. In contrast, other inputs have statistically significant effects. Based on the magnitude of the coefficients, the average within breeding herd prevalence is most sensitive to the apparent within breeding herd prevalence.

In order to evaluate robustness of rankings, the 95 percent confidence intervals are estimated for quantitative inputs. Comparison of the 95 percent confidence intervals for the apparent within breeding herd prevalence and the test sensitivity indicates that the rankings for these inputs are clear, because there is no overlap for the estimated confidence intervals.

If instead of the coefficient estimates, the magnitude of the F values is used as a criterion for ranking the inputs, the apparent within breeding herd prevalence, the study, and the test sensitivity will be ranked first to third, respectively.

Figure 6-4 illustrates the effect of the study on the fitted linear model. The linear relation between the output and the apparent breeding herd prevalence is plotted for each study level. The output is somewhat more sensitive to the apparent within breeding herd prevalence than to the choice of the study level. For example, for an apparent within breeding herd prevalence of 0.3, the output varies from approximately 45 to 95 percent depending upon the choice of study, or a range of approximately 50 percentage points. In contrast, for a given choice of study, such as Hancock/CFSAN (2001), the output varies between approximately 20 and 100 percent with respect to the apparent within breeding herd prevalence, or a range of approximately 80 percentage points. Thus, the typical range of variation in the average within breeding herd prevalence is much wider with respect to values for the apparent within breeding herd prevalence than it is with respect to the choice of study.

The $R^2$ for the linear regression model fitted to the dataset is 0.84 implying that the linear assumption for the functional relation between the output and inputs is substantially valid.

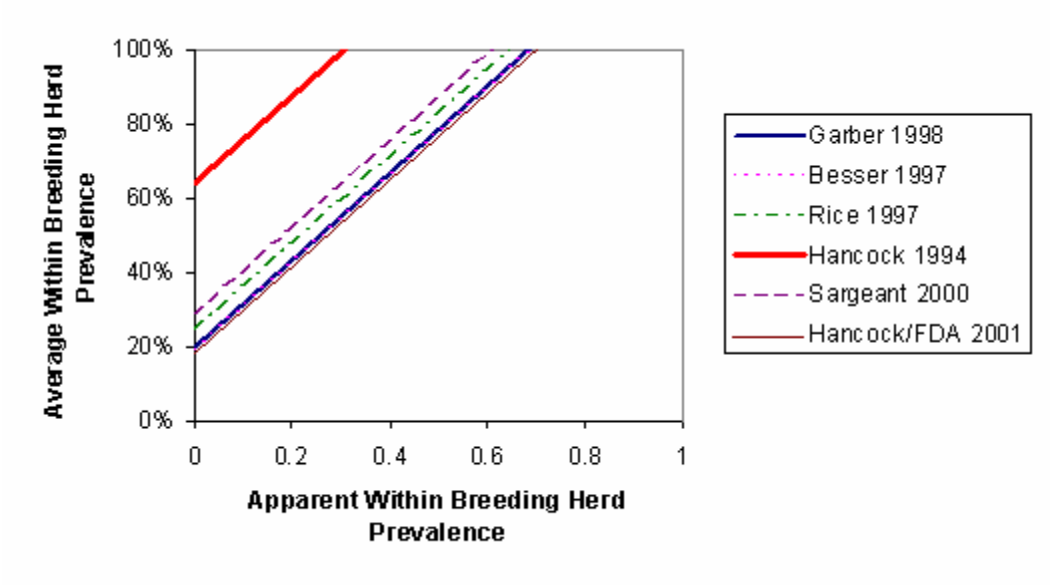



Figure 6-4. Regression Lines for Different Study Levels in the Within Breeding Herd Prevalence Part in Summer.

## 6.2 Regression Analysis in the Slaughter Module

The slaughter module is discussed in Section 3.2.2. Inputs and corresponding distributions in the slaughter module are summarized in Table 3-10. Three different types of probabilistic analysis were performed for this module, as described in Section 3.3.2: (1) one-dimensional simulation of variability based upon mean values of uncertain inputs; (2) two-dimensional simulation of variability for each realization of uncertainty; and (3) one-dimensional simulation of both variability and uncertainty co-mingled. In this section, the results of regression analysis for each of these three types of simulations are given. The case study scenario for the slaughter module is focused upon steers and heifers in the high prevalence season.

In Section 6.2.1, the results of regression analysis are presented based upon simulation of variability only. In Section 6.2.2, results are presented based upon the two-dimensional simulation of variability for different realizations of uncertainty. Results for the co-mingled one-dimensional simulation of both variability and uncertainty are given in Section 6.2.3. Section 6.2.4 compares the results from Sections 6.2.1 to 6.2.3.

### 6.2.1 Variability Only

This section presents the results of regression analysis applied to a one dimensional probabilistic simulation in which variability is only considered for mean uncertainty, based upon the case study scenario described in Section 3.3.2. The results of the regression analysis are given in Table 6-5. Inputs in Table 6-5 are ranked based on the magnitude of regression coefficients. These rankings are only presented for statistically significant inputs with Pr>F less than 0.05.

The chilling effect is the top ranked input with a coefficient of 0.3 and a 95 percent confidence interval of 0.29 to 0.31. The contaminated $cm^2$ is the second ranked input with a coefficient of 0.16 and a 95 percent confidence interval of 0.11 to 0.20. Because these two intervals do not overlap, the ranking of the first input is considered to be unambiguous. The third ranked input, washing effect ($W_{eff}$) has a coefficient of 0.13 and a confidence interval of 0.10 to 0.15. Because the confidence intervals for the second and third inputs overlap, the ranks are ambiguous for these two inputs. There are two inputs ranked fourth because they have coefficients with the same magnitude of 0.11. These inputs are the number of organisms ($N_{org}$) and the total number of contaminated animals (TNC). The magnitudes of the confidence intervals for these two inputs are 0.09 to 0.13 and 0.05 to 0.17, respectively. Both of these confidence intervals overlap with those of the third ranked input. Moreover, the interval for TNC overlaps with that of the second ranked input. Thus, the two fourth ranked inputs may be of comparable importance to the third ranked input, and one of the fourth ranked inputs could be of comparable importance to the second ranked input.

The confidence interval for the coefficient of the fifth ranked input, Trim/Vacuum/Wash efficiency, overlaps with that of the fourth ranked TNC input and of the sixth ranked input. The confidence interval for the coefficient of the sixth ranked input overlaps with that of the seventh ranked input. The other four inputs did not have statistically significant coefficients. Therefore, the latter four were deemed to be insensitive and are not ranked.

Because the coefficient confidence intervals typically overlap among closely ranked inputs for the third through seventh inputs, it is difficult to separate these inputs into groups that have clearly different importance. In general, it appears that the first rank is unambiguous. The second through fourth inputs, which include a total of four variables because two were tied for fourth, have overlapping confidence intervals as described above such that these four inputs may

Table 6-5. Regression Analysis Results for the Steer and Heifer Combo Bin Contamination in Summer for the Variability Analysis ($R^2 = 0.12$)

| Variable | Coefficient | 95% CI [(a)] | F Value | Pr>F | Rank |
|---|---|---|---|---|---|
| Total Number of Combo Bin for Each Carcass (TNCB) | $-7\times10^{-3}$ | (-0.020, -0.006) | 1 | 0.3 | --- |
| Total Number of Infected Animals (TNI) | $-4\times10^{-4}$ | (-0.03,0.03) | 0.1 | 0.9 | --- |
| Total Number of Contaminated Animals (TNC) | -0.11 | (-0.17, -0.05) | 15 | 0.0001 | 4 |
| Probability of Positive Cases at both Steps of Dehiding and Evisceration ($P_{both}$) | $-5\times10^{-3}$ | (-0.020,0.009) | 0.5 | 0.5 | --- |
| Number of Positive Cases at both Steps of Dehiding and Evisceration ($N_{both}$) | -0.04 | (-0.06, -0.02) | 22 | <0.0001 | 6 |
| Number of Positive Cases at Evisceration (NPE) | 0.01 | (0,0.02) | 2 | 0.2 | --- |
| Chilling Effect ($CH_{eff}$) | 0.3 | (0.29,0.31) | 1,640 | <0.0001 | 1 |
| Number of Organisms ($N_{org}$) | 0.11 | (0.09,0.13) | 155 | <0.0001 | 4 |
| Trim/Vacuum/Washing Efficiency (TVW) | -0.05 | (-0.07, -0.03) | 28 | <0.0001 | 5 |
| Evisceration Organisms Added ($N_{evisc}$) | 0.02 | (0.0,0.03) | 5 | 0.03 | 7 |
| Washing Effect ($W_{eff}$) | 0.13 | (0.10,0.15) | 160 | <0.0001 | 3 |
| Contaminated $cm^2$ (CCM) | 0.16 | (0.11,0.20) | 45 | <0.0001 | 2 |

(a) CI = Confidence Interval for the coefficient

be of comparable importance. The inputs ranked fifth, sixth, and seventh have substantially smaller average coefficients than those of the second through fourth ranked inputs. Although there is some overlap among the fourth ranked input for TNC with respect to the fifth and sixth ranked inputs, it is reasonable to consider the fifth through seventh ranked inputs as generally less important than the group that includes the second through fourth ranked inputs. Thus, the bottom line is that there appear to be four groups of inputs: (1) the top ranked input; (2) inputs of comparable but only moderate importance that are ranked second through fourth; and (3) inputs of comparable but only minor importance that are ranked fifth through seventh; and (4) four inputs that are of insignificant importance.

If F values were used instead of the magnitude of the coefficients as a basis for ranking, the results would be somewhat similar but not identical. The top ranked input is the same in

either case, since the chilling effect has an F value of 1,640, which is approximately an order-of-magnitude larger than the next largest F value. However, the input ranked second based upon coefficients does not have the second largest F value. The input with the second largest F value is the washing effect, which was ranked third based upon coefficients. Although the washing effect has a smaller coefficient than the contaminated $cm^2$, it has a narrower confidence interval for the coefficient. Thus, if F values were used as a basis for ranking, the following groups would emerge: (1) chilling effect clearly has the largest F value; (2) the washing effect and the number of organisms have comparable F values of approximately 160; (3) the contaminated $cm^2$ has an F value of 45 that is substantially smaller than that of the second group and substantially larger than that of the third group; (4) the Trim/Vacuum/Wash efficiency and the number of positive cases at both steps of dehiding and evisceration have comparable F values; (5) the evisceration organisms added has a small but statistically significant F value; and (6) four inputs have F values of 2 or less and are considered to be statistically insignificant. Rankings based upon the F value appear to place more importance on the confidence with which the regression coefficient is known, as opposed to only the magnitude of the coefficient.

The $R^2$ for the linear regression model fitted to the dataset equals 0.12. Thus, ranking based on the magnitude of the linear regression coefficients may not be reliable. Section 11.1.2 presents the comparison of the results based on the standardized linear regression with other methods that do not assume specific functional relationships, such as ANOVA and CART. The results in Table 11-5 indicate that rankings based on the linear regression analysis are substantially comparable to that of the other methods with respect to the selection of key inputs. Thus, even though the $R^2$ value in this case is low, the ranking of the inputs is similar to that obtained with other methods.

### 6.2.2 Two-Dimensional Simulation of Variability for Different Uncertainty Realizations

This case study is based upon a two-dimensional simulation of variability with respect to different realizations of uncertainty. The simulation has a total sample size of 65,000, based upon 650 variability iterations for each of 100 uncertainty iterations. The objective was to identify the most sensitive inputs with respect to variability. Therefore, regression analysis was applied for each of the 100 uncertainty iterations, resulting in 100 alternative rankings of the key inputs.

Table 6-6. Summary of the Regression Analysis Results for Two-Dimensional Variability Simulation for Different Uncertainty Realizations (Mean $R^2 = 0.38$)

| Variable [1] | Mean Coefficient | 95% Probability Range of Coefficients | Frequency [2] | Mean Rank | Range of Rank |
|---|---|---|---|---|---|
| TNCB | -0.003 | (-0.05, 0.04) | 5 | 10.6 | 7 – 12 |
| TNI | -0.02 | (-0.22, 0.11) | 19 | 8.3 | 3 – 12 |
| TNC | -0.09 | (-1.26, 0.68) | 51 | 4.5 | 1 – 12 |
| $P_{both}$ | -0.02 | (-0.11, 0.04) | 18 | 9.7 | 4 – 12 |
| $N_{both}$ | 0.004 | (-0.24, 0.23) | 45 | 8.1 | 2 – 12 |
| NPE | -0.03 | (-1.37, 0.44) | 43 | 6.8 | 1 – 12 |
| $CH_{eff}$ | 0.62 | (0.04, 1.02) | 90 | 2.2 | 1 – 8 |
| $N_{org}$ | 0.26 | (-0.14, 0.77) | 92 | 4.4 | 1 – 11 |
| TVW | 0.09 | (-0.8, 0.28) | 56 | 6.3 | 1 – 12 |
| $N_{evisc}$ | 0.09 | (-0.45, 0.95) | 44 | 6.5 | 1 – 12 |
| $W_{eff}$ | 0.057 | (-0.38, 0.55) | 64 | 6.2 | 1 – 11 |
| CCM | 0.158 | (-0.44, 1.20) | 59 | 4.3 | 1 - 9 |

(1) See Table 6-5 for definition of variable names.
(2) The percentage of the 100 uncertainty simulations for which the coefficient was statistically significant.

The inputs included in regression analysis for the two-dimensional simulation were the same as those for the one-dimensional simulation of variability only as listed in Table 6-5. The results of the 100 regression analyses are summarized in Table 6-6. The table includes the mean coefficient estimate of each input, 95 percent probability range for each coefficient, and the range of ranks for each input in 100 uncertainty realizations. The percentage of the 100 simulations that produced a statistically significant coefficient is also quantified. Furthermore, the mean rank for a given input is specified.

Mean ranks over 100 uncertainty realizations in Table 6-6 indicate that the chilling effect is the most important input. There is 90 percent probability that the chilling effect is identified as a statistically significant input in uncertainty realizations and it has a mean rank of 2.2. The mean ranks for the number of organisms, total number of contaminated animals, and number of contaminated $cm^2$ are estimated as 4.4, 4.5, and 4.3 indicating that on average the output shows approximately the same sensitivity to these inputs. For these inputs the probabilities of being statistically significant in 100 uncertainty realizations are 92, 51, and 59 percent, respectively. The Trim/Vacuum/Wash efficiency, the number of organisms added due to evisceration, washing efficiency, and number of positive cases at evisceration have mean ranks of 6.3, 6.5, 6.2, and 6.8 with probabilities of being statistically significant of 56, 44, 64 and 43

percent, respectively. Hence, the output shows approximately the same sensitivity to these inputs. The output shows on average the same sensitivity to the total number of infected animals and the number of contaminated animals at both steps which have mean ranks of 8.3 and 8.1, respectively. The output presents the lowest sensitivity to the probability of positive cases at both steps of dehiding and evisceration and the number of combo bins to which each animal contributes. These latter two inputs have mean ranks of 9.7 and 10.6.

In order to visualize the results of the sensitivity analysis, the complementary cumulative distribution function (CCDF) of the rank is given for each input based upon the 100 uncertainty realizations. Figure 6-5 displays the CCDFs for four inputs that have the highest average ranks among all of the inputs included in the analysis. These inputs are chilling effect ($CH_{eff}$), total number of contaminated animals (TNC), number of contaminated $cm^2$ (CCM), and number of organisms on the carcass surface ($N_{org}$). The CCDF for the chilling effect indicates that for approximately 45 percent of the simulations, the rank was worse than one, which implies that the rank was equal to one for 55 percent of the simulations. Furthermore, the chilling effect was ranked five or higher for 90 percent of the simulations. In contrast, total number of contaminated animals, number of contaminated $cm^2$, and number of organisms on the carcass surface were ranked first for 5, 11, and 14 percent of the simulations, respectively, and were ranked fifth or higher for 59, 62, and 69 percent of the simulations.

When comparing the CCDFs of Figure 6-5, it is apparent that the chilling effect tends to have a higher rank than the other inputs. Furthermore, because the probability that the chilling effect has a rank of five or higher is high, the identification of the chilling effect as one of the most important inputs is robust to uncertainty. In contrast, the total number of contaminated animals, number of contaminated $cm^2$, and number of organisms on the carcass surface have 41, 38, and 31 percent probability, respectively, of having a rank worse than five. Thus, although these three inputs typically have a similar importance to each other, they are typically less important than the chilling effect.

Figure 6-6 displays the CCDFs for four inputs that have the middle average ranks between five and eight among all of the inputs included in the analysis. These inputs are Trim/Vacuum/Washing efficiency (TVW), washing efficiency ($W_{eff}$), number of positive cases at evisceration (NPE), number of organisms added due to evisceration ($N_{evisc}$). The CCDF for the number of organisms added due to evisceration indicates that for approximately 88 percent of the

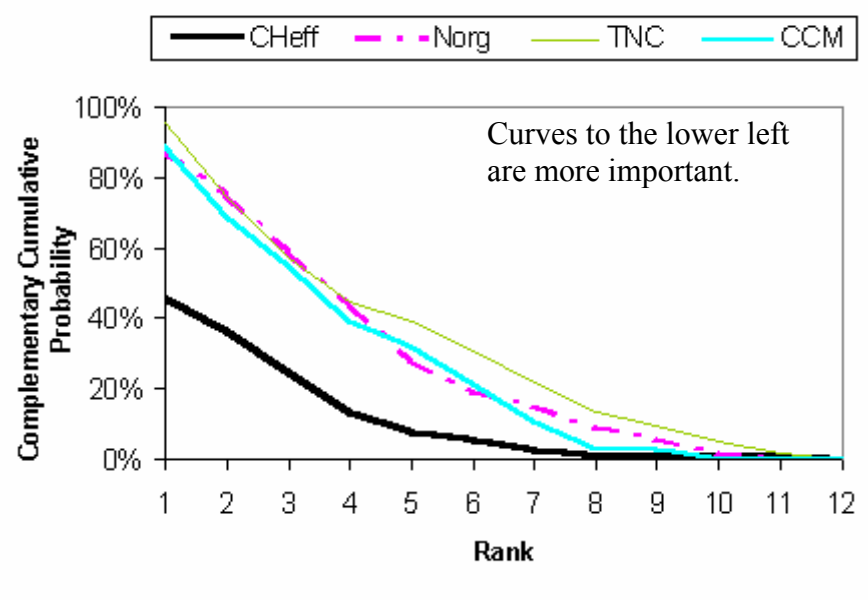



Figure 6-5. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected Inputs Based Upon Regression Analysis: Chilling Effect ($CH_{eff}$); Total Number of Contaminated Animals (TNC); Number of Contaminated $cm^2$ (CCM); and Number of Organisms on the Carcass Surface ($N_{org}$).

simulations, the rank was worse than one, which implies that the rank was equal to one for 12 percent of the simulations. In contrast, for other inputs, the probability of the rank being worse than one is approximately 95 percent. The mean ranks of these inputs are between 6.2 and 6.8. The CCDF graphs indicate that the probability of having ranks worse than eight for these four inputs ranges from 19 to 36. When comparing the CCDFs, it is apparent that the identification of the rank of these inputs is not robust to uncertainty. The rank for these inputs varies between one and twelve based on different uncertainty realizations. Hence, there is ambiguity regarding the rank of each input as a function of uncertainty in the model inputs.

The least important group of inputs is depicted in Figure 6-7. These inputs include the total number of combo bins to which each animal contributes (TNCB), probability of positive cases at both steps of dehiding and evisceration ($P_{both}$), the total number of positive cases at both steps of dehiding and evisceration ($N_{both}$), and the total number of infected animals (TNI). These inputs have a probability ranging from 78 to 100 percent of having a rank worse than five, and their average ranks range from 8.1 to 10.6. $N_{both}$ and TNI have similar CCDF distributions. The similarity of these distributions implies that these two inputs are of comparable importance. All four of these inputs are typically ranked seven or worse for approximately 70 to 95 percent of the

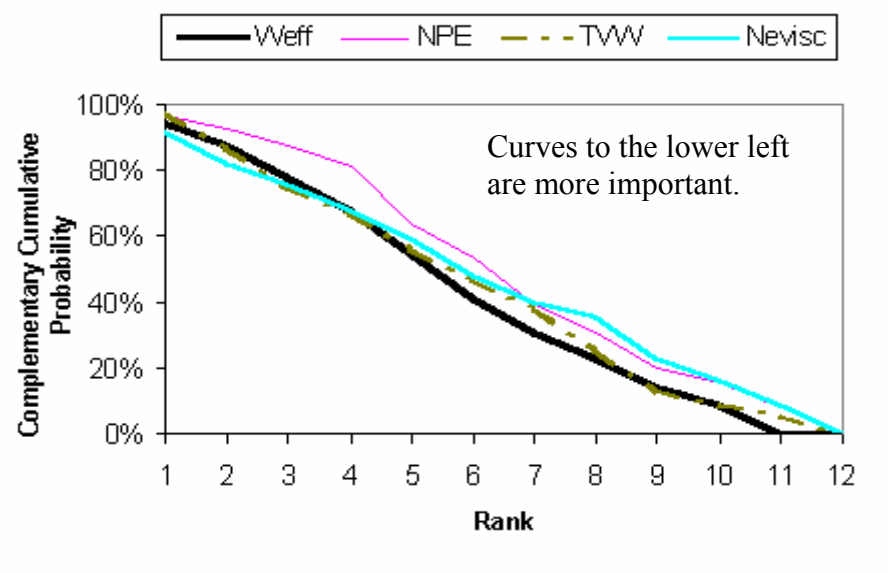



Figure 6-6. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected Inputs Based Upon Regression Analysis: Washing Efficiency ($W_{eff}$); Number of Positive Cases at Evisceration (NPE); Number of *E. coli* Organisms Added Due to Evisceration ($N_{evisc}$); and Trim/Vacuum/Washing Efficiency (TVW).

uncertainty realizations. Moreover, there are a few uncertainty iterations for which these inputs have ranks as high as three or four.

A comparison of the three figures helps gain insight into how the inputs should be grouped with respect to their importance. Chilling effect has the highest probability of a rank of one and the CCDF for this input is clearly different than those of $N_{org}$, TNC, and CCM. These latter three have similar CCDFs and therefore are of comparable but secondary importance compared to the chilling effect. The four inputs in Figure 6-6 are of comparable importance because their CCDFs are similar to each other. Furthermore, the group of inputs in Figure 6-6 is generally less important than the group of three inputs in Figure 6-5 that are of secondary importance. Thus, the four inputs in Figure 6-6 are judged to comprise a group representing the third most important set of inputs. The two inputs, $N_{both}$ and TNI, have similar CCDFs and tend to have worse ranks than the third most important set of inputs. Therefore, these two inputs are judged to comprise a group of fourth importance. Finally, the remaining two inputs shown in Figure 6-5 that have the lowest average ranks are judged to be the least important.

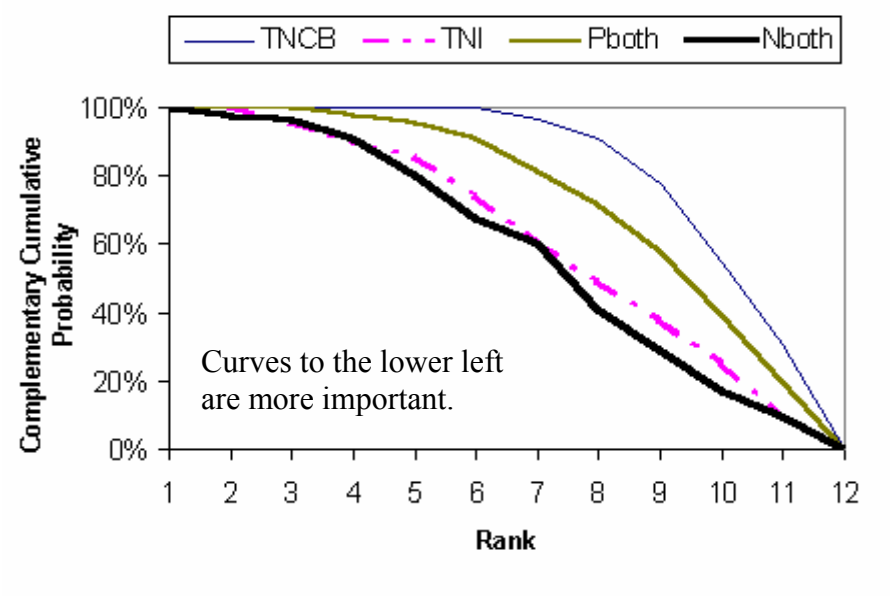



Figure 6-7. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected Inputs Based Upon Regression Analysis: Probability of Positive Cases at both Steps of Dehiding and Evisceration ($P_{both}$); Total Number of Combo Bins (TNCB); Total Number of Infected Animals (TNI); and Number of Positive Cases at both Steps of Dehiding and Evisceration ($N_{both}$).

### 6.2.3 One-Dimensional Simulation of Variability and Uncertainty

This section presents the results of regression analysis applied to a one dimensional probabilistic simulation in which variability and uncertainty are co-mingled, based upon the case study scenario described in Section 3.3.2.

Table 6-7 summarizes the results of application of regression analysis to the slaughter module for the co-mingled simulation of variability and uncertainty. The inputs in Table 6-7 are ranked based on the magnitude of regression coefficients. Rankings are presented for statistically significant inputs with Pr>F less than 0.05. The F values in Table 6-7 indicate that the following inputs are not significant: total number of combo bins to which each carcass contributes; the total number of infected animals; the number of positive cases at both steps of dehiding and evisceration; and the number of positive cases at evisceration.

Based upon the magnitude of the coefficients for the statistically significant inputs, the chilling effect, the number of organisms on contaminated carcasses, and the washing effect are the top three sensitive inputs. In order to evaluate the robustness of the estimated rankings, the 95 percent confidence intervals are estimated for each coefficient. Estimated confidence intervals for regression coefficients indicate that the rankings for the top three inputs are robust. For

Table 6-7. The Regression Analysis Results for Steer and Heifer Combo Bin Contamination in Summer Based Upon One-Dimensional Co-Mingled Variability and Uncertainty Simulation ($R^2$ = 0.10)

| Variable | Coefficient | 95% CI [(a)] | F Value | Pr>F | Rank |
|---|---|---|---|---|---|
| Total number of combo bin for each carcass | $-6 \times 10^{-4}$ | $(-8,7)\times10^{-3}$ | 0.02 | 0.9 | --- |
| Total number of infected animals | -0.01 | (-0.02,0) | 3 | 0.1 | --- |
| Total number of contaminated animals | 0.05 | (0.03,0.07) | 45 | <0.0001 | 4 |
| Probability of positive cases at 2 steps | -0.02 | (-0.03, -0.01) | 15 | 0.0001 | 6 |
| Number of positive cases at 2 steps | $4\times10^{-3}$ | $(-5,14)\times10^{-3}$ | 0.8 | 0.4 | --- |
| Number of positive cases at evisceration | $-4\times10^{-3}$ | $(-10,5)\times10^{-3}$ | 0.8 | 0.4 | --- |
| Chilling effect | 0.2 | (0.19,0.21) | 2,800 | <0.0001 | 1 |
| Number of organisms | 0.13 | (0.12,0.14) | 920 | <0.0001 | 2 |
| Trim vacuum washing efficiency | -0.04 | (-0.05, -0.03) | 65 | <0.0001 | 5 |
| Evisceration organisms added | 0.05 | (0.04,0.06) | 170 | <0.0001 | 4 |
| Washing effect | 0.09 | (0.08,0.10) | 425 | <0.0001 | 3 |
| Contaminated $cm^2$ | 0.02 | (0.01,0.03) | 16 | <0.0001 | 6 |

(a) CI = Confidence Interval for the coefficient

example, the confidence intervals for the chilling effect, the number of organisms on contaminated carcasses, and the washing effect do not overlap indicating that their ranks are robust.

There is overlap in the magnitudes of the 95 percent confidence intervals for the coefficients when comparing both the fourth and the fifth ranked inputs, and when comparing both of the sixth ranked inputs (two inputs had the same magnitude for the coefficient). Thus, there are essentially three groups of inputs. The most sensitive group includes the top three ranked inputs that are significantly different from each other in importance. The second group includes the fourth and fifth ranked inputs, which are of comparable importance because their coefficient confidence intervals overlap. The third group includes inputs ranked sixth that are of comparable minor yet statistically significant importance.

The use of F values instead of coefficients to gain insight into key inputs would lead to similar rankings. The F values not only take into account the magnitude of the coefficient, but also consider the amount of error corresponding to each coefficient. The top three inputs identified above has the same rankings when F values are used instead. The same three inputs are included in the second group of inputs, and the same two inputs are included in the third group. Within the second group, there is some difference in rankings based upon the F values compared to those based upon the average coefficients. Because the average coefficients were not significantly different from each other within this group, the differences in ranking among them when comparing the two approaches are deemed to be insignificant.

The $R^2$ for the linear regression model fitted to the dataset is 0.10 indicating that the ranks based on the magnitude of the linear regression coefficients may not be reliable. Results of the analysis in this part are compared in Section 11.1.2 to that of the other methods such as ANOVA and CART that do not impose any specific functional relationships. Table 11-7 indicates that rankings based on the linear regression analysis are substantially comparable to that of the other methods with respect to the selection of key inputs. Therefore, although the $R^2$ value is low the results in this case are similar to those of other methods.

### 6.2.4 Summary and Comparison of the Results of Regression Analysis in the Slaughter Module

In Sections 6.2.1 to 6.2.3 regression analysis was applied to three datasets considering variability only, variability for different uncertainty realizations, and co-mingled variability and uncertainty in inputs. In this section rankings based on these analyses are summarized and compared. Table 6-8 gives the ranks for each input based on analyses in Sections 6.2.1 to 6.2.3.

The key similarities among the three probabilistic simulations were with respect to the identification of the most important input and the least important inputs. All three probabilistic analysis methods resulted in identification of chilling effect as clearly the most important input. All three approaches resulted in identification of the total number of combo bins to which each carcass contributes, the total number of infected animals, the probability of positive cases at both steps of dehiding and evisceration, the number of positive cases at both steps dehiding and evisceration, and the number of positive cases at evisceration as among the least important inputs. Inputs that were of moderate importance based upon each of three methods were similar. For example, the number of organisms added, total number of contaminated animals, and

Table 6-8. Summary of the Regression Analysis Results Based on Variability Only, Variability for Different Uncertainty Realizations, and Co-mingled Variability and Uncertainty Analyses

| Variable | Ranks | | |
|---|---|---|---|
| | Analysis 1(1) | Analysis 2(2) | Analysis 3(3) |
| Total Number of Combo Bins for Each Carcass | --- | 10.6 | --- |
| Total Number of Infected Animals | --- | 8.3 | --- |
| Total Number of Contaminated Animals | 4 | 4.5 | 4 |
| Probability of Positive Cases at both Steps of Dehiding and Evisceration | --- | 9.7 | 6 |
| Number of Positive Cases at both Steps of Dehiding and Evisceration | 6 | 8.1 | --- |
| Number of Positive Cases at Evisceration | --- | 6.8 | --- |
| Chilling Effect | 1 | 2.2 | 1 |
| Number of Organisms | 4 | 4.4 | 2 |
| Trim/Vacuum/Washing Efficiency | 5 | 6.3 | 5 |
| Evisceration Organisms Added | 7 | 6.5 | 4 |
| Washing Effect | 3 | 6.2 | 3 |
| Contaminated $cm^2$ | 2 | 4.3 | 6 |

(1) Ranks based on the variability only analysis.
(2) Mean ranks based on the variability for different uncertainty realizations analysis.
(3) Ranks based on the one-dimensional co-mingled variability and uncertainty analysis.

washing effect were typically in the upper or middle tier of inputs for all three approaches.

There were some inputs for which the rankings appear to be different based upon the three simulation methods. For example, the contaminated $cm^2$ of meat trims is ranked second based upon variability only, $4.3^{rd}$ based upon the two-dimensional simulation, and $6^{th}$ based upon the one-dimensional simulation of both variability and uncertainty. For the variability only case, this input was not significantly different from one of the fourth ranked inputs. For the two-dimensional case, this input was not substantially different in importance compared to two other inputs. For the co-mingled one-dimensional simulation of both variability and uncertainty, this input was clearly in the least important statistically significant group. Thus, the results for the variability only and the two dimensional simulations are approximately similar, but both of these differ from the results of the one-dimensional simulation of both variability and uncertainty.

Although some of the middle ranked inputs had different rankings when comparing the three simulation methods, the most important finding is that the ranking of the top input and of the least important inputs was essentially the same for all three approaches. This implies that

any of the three approaches could be used alone in order to make a distinction between the top input and the unimportant inputs. The differences in rankings for the significant inputs of moderate importance suggest that the rankings are sensitive to the actual ranges of values used in the probabilistic simulations. The co-mingled simulation of both variability and uncertainty is expected to produce the widest ranges of values within a single simulation compared to the simulation of only variability. Thus, it is expected that the results of these two approaches should differ. The two dimensional approach distinguishes between variability and uncertainty. Thus, although the range of values for each input over the course of the entire simulation is similar to that for the one dimensional approach in which variability and uncertainty are co-mingled, the range of values for any given realization of variability will typically be comparable to that of the one dimensional simulation of variability only. Thus, the analysis based upon two-dimensional simulation is expected to produce results somewhat different than those from the other two methods.

## 6.3 Regression Analysis in the Preparation Module

In the preparation module regression analysis was applied to two parts, including growth estimation and serving contamination parts. The results of the analyses for these two parts are presented in Sections 6.3.1 and 6.3.5, respectively. In order to compare alternative regression-based approaches several methods are applied to the growth estimation part. These methods include Pearson sample correlation coefficients, Spearman rank correlation coefficients, and rank regression. Results are presented in Sections 6.3.2 and 6.3.3 for correlation coefficients and rank regression, respectively. Section 6.3.4 presents the comparison of the results based on the regression analysis, correlation coefficients, and rank regression.

Regression analysis was not applied to the cooking effect part. As given in Table 3-13, this part has only one quantitative input, with other inputs being qualitative. Although shown in earlier sections of this chapter that the F values can be used to make inferences regarding the sensitivity of qualitative inputs, for situations in which qualitative inputs predominate a judgment was made that other methods, such as ANOVA, are better suited than regression analysis. The application of ANOVA to the cooking effect part is given in Section 5.4.2.

### 6.3.1 Regression Analysis in the Growth Estimation Part

The growth estimation part is discussed in Section 3.2.3. Three different types of probabilistic analysis were performed for this part, as described in Section 3.3.3: (1) one-

Table 6-9. Regression Analysis Results for the Growth Estimation Part Based Upon Variability only ($R^2 = 0.51$)

| Variable | Coefficient | 95% CI [(a)] | F Value | Pr>F | Rank |
|---|---|---|---|---|---|
| Storage Temperature, Stage 1 ($Temp_1$) | 0.32 | (0.31,0.32) | 5,500 | <0.0001 | 3 |
| Storage Temperature, Stage 2 ($Temp_2$) | $3\times10^{-3}$ | $(-7,13)\times10^{-3}$ | 0.4 | 0.6 | --- |
| Storage Temperature, Stage 3 ($Temp_3$) | 0.59 | (0.58,0.60) | 15,500 | <0.0001 | 1 |
| Storage Time, Stage 1 ($Time_1$) | 0.27 | (0.26,0.27) | 9,400 | <0.0001 | 4 |
| Storage Time, Stage 2 ($Time_2$) | $7\times10^{-3}$ | $(2,13)\times10^{-3}$ | 8 | 0.1 | --- |
| Storage Time, Stage 3 ($Time_3$) | 0.34 | 0.34[(b)] | 15,200 | <0.0001 | 2 |
| Maximum Density (MD) | 0.012 | (0.007,0.02) | 21 | <0.0001 | 7 |
| Lag Period, Stage 1 ($LP_1$) | -0.012 | (-0.019, 0.01) | 11 | 0.0005 | 7 |
| Lag Period, Stage 2 ($LP_2$) | $-1\times10^{-4}$ | (-0.01,0.01) | 0.0 | 0.3 | --- |
| Lag Period, Stage 3 ($LP_3$) | $-1\times10^{-3}$ | (-0.01,0.01) | 1.1 | 0.4 | --- |
| Generation Time, Stage 1 ($GT_1$) | 0.08 | (0.07,0.09) | 390 | <0.0001 | 6 |
| Generation Time, Stage 2 ($GT_2$) | $-4\times10^{-3}$ | (-0.014,0.01) | 0.5 | 0.7 | --- |
| Generation Time, Stage 3 ($GT_3$) | 0.11 | (0.10,0.12) | 530 | <0.0001 | 5 |

(a) CI = Confidence Interval for the coefficient
(b) 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.

dimensional simulation of variability based upon mean values of uncertain inputs; (2) two-dimensional simulation of variability for each realization of uncertainty; and (3) one-dimensional simulation of both variability and uncertainty co-mingled.

In the next section, the results of regression analysis are presented based upon simulation of variability only. In Section 6.3.1.2, results are presented based upon the two-dimensional simulation of variability for different realizations of uncertainty. Results for the co-mingled one-dimensional simulation of both variability and uncertainty are given in Section 6.3.1.3. Section 6.3.1.4 compares the results from Sections 6.3.1.1 to 6.3.1.3.

6.3.1.1 Variability Only

This section presents the results of regression analysis applied to a one-dimensional probabilistic simulation in which variability is only considered for mean uncertainty, based upon the case study scenario described in Section 3.3.3.

Table 6-9 summarizes the results of application of regression analysis to the growth estimation part for the simulation of variability only. The inputs are ranked based on the

magnitude of regression coefficients. Rankings are presented for statistically significant inputs with Pr>F less than 0.05. The F values indicate that there are no statistically significant effects for inputs the storage temperature, the storage time, lag period, and the generation time at stage 2 and the lag period at stage 3.

The rankings of the inputs based upon the coefficient values are generally unambiguous. The top ranked input 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. There are two inputs with ranked seventh with equal coefficients. The confidence intervals for these inputs overlap. However, both of these inputs have coefficients that are substantially smaller than all of the other inputs. Therefore, these two inputs are of little importance compared to the other ranked inputs.

Because the confidence intervals of the coefficients do not overlap in most cases, other than for the two seventh ranked inputs, there are few groups of inputs of similar importance. The inputs with ranks of four or higher have coefficients greater than 0.25. The fifth and sixth ranked inputs have coefficients between 0.08 and 0.11.

If F values were used instead of coefficients as a basis for ranking, the rankings would be similar. With the exception of the third ranked inputs for which the F value is substantially smaller than the fourth ranked input, the rankings based upon F values would be the same as those based upon the regression coefficients.

The $R^2$ for the linear regression model fitted to the dataset is 0.51. Although the $R^2$ value is not very high, it is still in an acceptable range indicating that ranking based on the magnitude of the linear regression coefficients may be reliable.

#### 6.3.1.2 Two-Dimensional Simulation of Variability for Different Uncertainty Realizations

The application of regression analysis to a two-dimensional simulation in which variability is simulated for each different realization of uncertainty involves sensitivity analysis for each of the uncertainty iterations. In this case, for example, there are 100 uncertainty iterations. Within each uncertainty iteration, 650 samples were generated to represent variability in each input. Thus, regression analysis was applied 100 times.

Table 6-10. Summary of the Regression Analysis Results for Two-Dimensional Variability Simulation for Different Uncertainty Realizations (Average $R^2 = 0.51$)

| Variable[(1)] | Mean Coefficient | 95% Probability Range of Coefficients | Frequency [(2)] | Mean Rank | Range of Rank |
|---|---|---|---|---|---|
| Temp1 | 0.31 | (0.01, 0.70) | 93 | 3.1 | 1 – 13 |
| Temp2 | 0.01 | (-0.09,0.10) | 3 | 8.5 | 4 – 13 |
| Temp3 | 0.54 | (0.29,0.83) | 100 | 1.4 | 1 – 3 |
| Time1 | 0.26 | (0.08,0.49) | 100 | 3.5 | 1 – 7 |
| Time2 | 0.00 | (-0.04,0.06) | 6 | 10.8 | 6 – 13 |
| Time3 | 0.34 | (0.13,0.49) | 99 | 2.7 | 1 – 4 |
| MD | 0.02 | (-0.04,0.07) | 11 | 10.1 | 5 – 13 |
| LP1 | -0.02 | (-0.11,0.07) | 21 | 8.9 | 3 – 13 |
| LP2 | -0.01 | (-0.09,0.06) | 5 | 9.7 | 5 – 13 |
| LP3 | 0.0 | (-0.09,0.08) | 7 | 9.7 | 5 – 13 |
| GT1 | 0.07 | (-0.02,0.22) | 40 | 7.2 | 3 – 13 |
| GT2 | 0.01 | (-0.08,0.09) | 4 | 9.0 | 5 – 13 |
| GT3 | 0.09 | (-0.03,0.22) | 45 | 6.4 | 3 - 13 |

(1) See Table 6-9 for definition of variable names
(2) The percentage of the 100 uncertainty simulations for which the coefficient was statistically significant.

The inputs included in regression analysis for the two-dimensional simulation were the same as those for the one-dimensional simulation of variability only as listed in Table 6-9. The results of the 100 analyses with regression analysis are summarized in Table 6-10. The table includes the mean coefficient estimate of each input, 95 percent probability range for each coefficient, and the range of ranks for each input in 100 uncertainty realizations. The percentage of the 100 simulations that produced a statistically significant coefficient is also quantified. Furthermore, the mean rank for a given input is specified.

Mean ranks over 100 uncertainty realizations in Table 6-10 indicate that the storage temperature at stage 3, which has a mean rank of 1.4, is the most important input. There is 100 percent probability that this input is identified statistically significant. Three inputs have approximately similar average rankings of 2.7 to 3.5 and similar range of rankings indicating that they are of comparable importance to each other but less important than the storage temperature at stage 3. These three inputs are the storage time at stage 3, the storage temperature at stage 1, and the storage time at stage 1. For these three inputs the probabilities of being statistically significant among all 100 uncertainty realizations are 99, 93, and 100 percent, respectively. Inputs related to the second stage such as the storage temperature, the storage time, lag period,

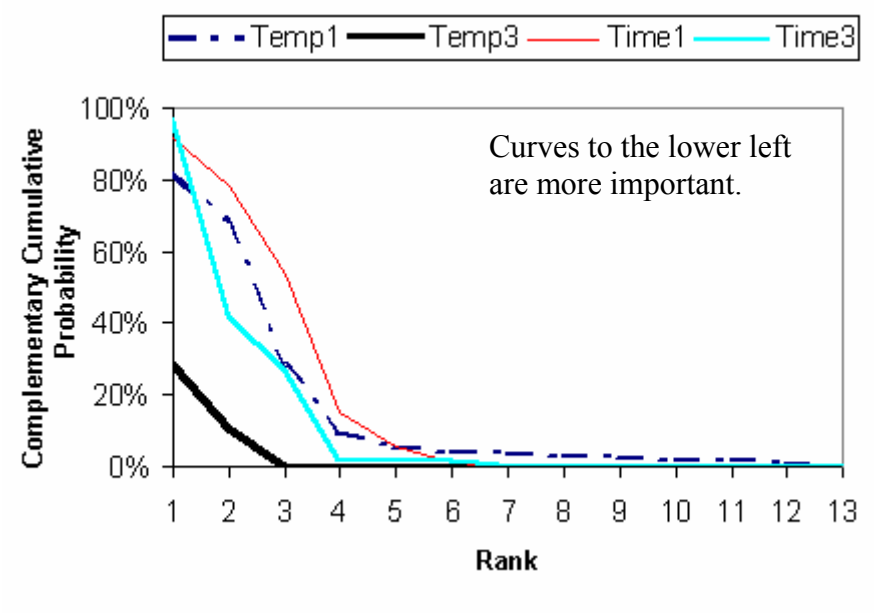



Figure 6-8. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected Inputs Based Upon Regression Analysis: Storage Temperature at Stages 1and 3 ($Temp_1$ and $Temp_3$); and Storage Time at Stages 1 and 3 ($Time_1$ and $Time_3$).

and generation time have least importance among all the inputs in 100 uncertainty realizations with average ranks ranging between 6.5 and 10.8 and probabilities of being statistically significant ranging between 3 to 6. Other inputs not already mentioned are typically of minor to little importance with average ranks of 6.4 to 10.1.

In order to visualize the results of the sensitivity analysis, the complementary cumulative distribution function (CCDF) of the rank is given for each input. Figure 6-8 displays the CCDFs for four inputs that have the highest average ranks among all of the inputs included in the analysis. These inputs are storage time at stage 3, storage temperature at stage 3, storage time at stage 1, and storage temperature at stage 1. The CCDF for the storage temperature at stage 3 indicates that for 28 percent of the simulations, the rank was worse than one, which implies that the rank was equal to one for 72 percent of the simulations.

Furthermore, the storage temperature at stage 3 was ranked three or higher for 100 percent of the simulations. In contrast, storage time at stages 1 and 3 and storage temperature at stage 1 were ranked first for 10, 3, and 20 percent of the simulations, respectively. These inputs were ranked fifth or higher for 93, 98, and 93 percent of the simulations, respectively. Thus, the storage temperature at stage 3 has the highest frequency of a rank of one. The other three inputs are of comparable importance and are each less important than the storage temperature at stage 1.

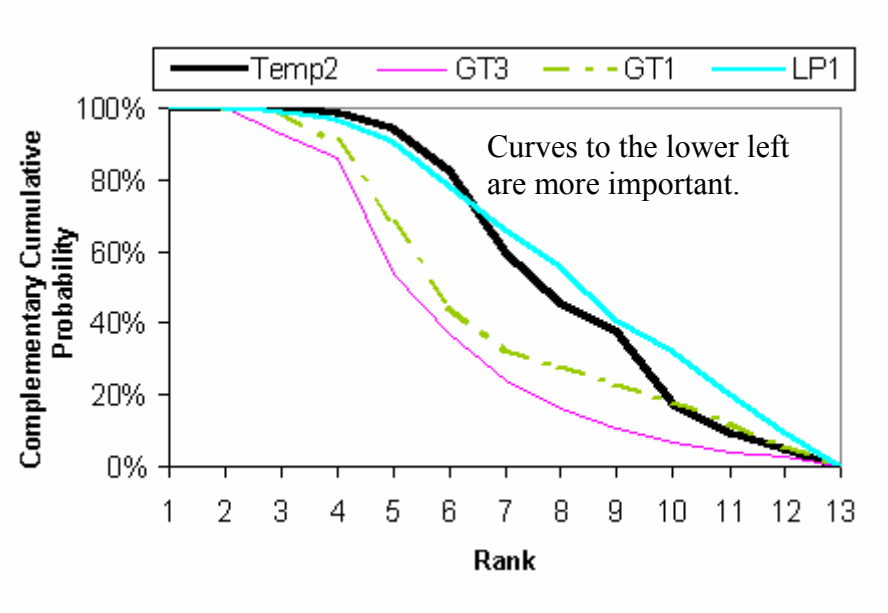



Figure 6-9. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected Inputs Based Upon Regression Analysis: Storage Temperature at Stage 2 ($Temp_2$); Generation Time at Stages 1 and 3 ($GT_1$ and $GT_3$); and Lag Period at Stage 1 ($LP_1$).

Furthermore, the storage temperature at stage 3 was ranked three or higher for 100 percent of the simulations. In contrast, storage time at stages 1 and 3 and storage temperature at stage 1 were ranked first for 10, 3, and 20 percent of the simulations, respectively. These inputs were ranked fifth or higher for 93, 98, and 93 percent of the simulations, respectively. Thus, the storage temperature at stage 3 has the highest frequency of a rank of one. The other three inputs are of comparable importance and are each less important than the storage temperature at stage 1.

Figure 6-9 displays the CCDFs for four inputs that have average ranks between six and nine among all of the inputs included in the analysis. These inputs are the storage temperature at stage 2 ($Temp_2$), generation times at stages 1 and 3 ($GT_1$ and $GT_3$), and lag period at stage ($LP_1$). The CCDF for these inputs indicate that for 90 percent or more of the simulations, the ranks for these inputs were worse than three. The probability that the inputs have ranks of worse than nine varies between 15 to 40 percent. The rank for these inputs varies based on different uncertainty realizations. Hence, there is ambiguity regarding the rank of each input as a function of uncertainty in the model inputs.

The least important group of inputs is depicted in Figure 6-10. These inputs include storage time at stage 2 ($Time_2$), maximum density (MD), generation time at stage 2 ($GT_2$), and lag periods at stages 2 and 3 ($LP_2$ and $LP_3$). These inputs have a probability ranging from 88 to

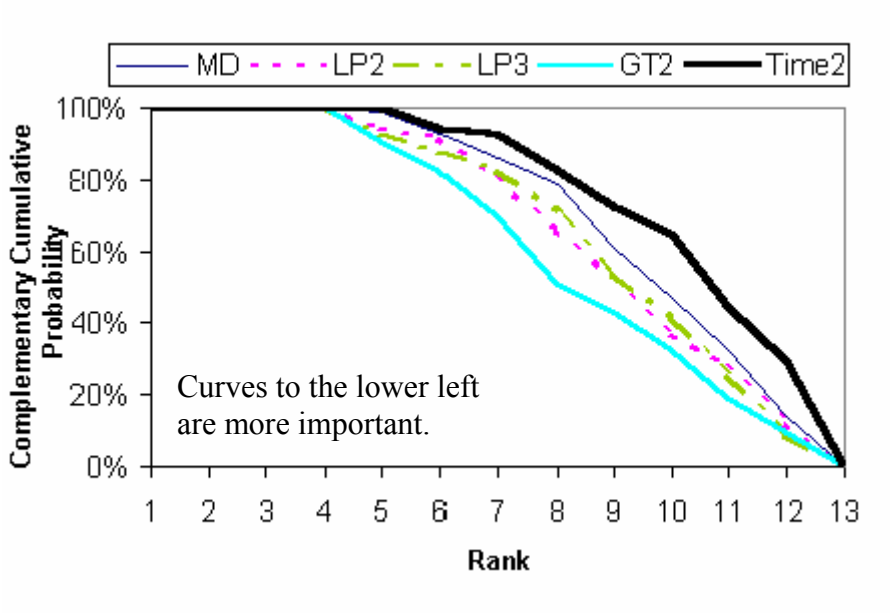



Figure 6-10. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected Inputs Based Upon Regression Analysis: Lag Period at Stages 2 and 3 ($LP_2$ and $LP_3$); Generation Time at Stage 2 ($GT_2$); Maximum Density (MD); and Storage Time at Stage 2 ($Time_2$).

100 percent of having a rank worse than five, and their average ranks range from 9 to 11. MD, $LP_2$ and $LP_3$ have similar CCDF distributions. The similarity of these distributions implies that these three inputs are of comparable importance. There is ambiguity regarding the rank of each input as a function of uncertainty in the model inputs. $Time_2$ can be identified as the least sensitive input based on the CCDF distribution. $Time_2$ has a rank worse than 10 with probability of 65 percent. Moreover, this input was statistically insignificant for 94 percent of the uncertainty realizations.

The results shown in the three figures indicate that there are approximately four groups of inputs. These groups include: (1) the most important input of storage temperature at stage 3; (2) inputs of secondary importance, including storage time at stages 1 and 3 and storage temperature at stage 1; (3) inputs of tertiary importance, including generation times at stage 1 and 3; and (4) inputs of minor or no importance, including storage temperature and time at stage 2, lag periods at stages 1, 2, and 3, generation time at stage 2, and maximum density.

Table 6-11. The Regression Analysis Results for the Growth Estimation Part Based Upon One-Dimensional Co-mingled Variability and Uncertainty Simulation ($R^2 = 0.50$)

| Variable | Coefficient | 95% CI [(a)] | F Value | Pr>F | Rank |
|---|---|---|---|---|---|
| Storage Temperature, Stage 1 | 0.32 | (0.31,0.32) | 5,900 | <0.0001 | 3 |
| Storage Temperature, Stage 2 | 0.01 | (0.0,0.02) | 5 | 0.03 | --- |
| Storage Temperature, Stage 3 | 0.55 | (0.54,0.56) | 13,700 | <0.0001 | 1 |
| Storage Time, Stage 1 | 0.27 | (0.27,0.28) | 10,000 | <0.0001 | 4 |
| Storage Time, Stage 2 | $6\times10^{-4}$ | $(-5,6)\times10^{-3}$ | 0.1 | 0.8 | --- |
| Storage Time, Stage 3 | 0.37 | (0.37,0.38) | 18,700 | <0.0001 | 2 |
| Maximum Density | 0.02 | (0.01,0.03) | 75 | <0.0001 | 7 |
| Lag Period, Stage 1 | -0.02 | (-0.03, -.01) | 35 | <0.0001 | 7 |
| Lag Period, Stage 2 | -0.01 | (-0.02,0.0) | 1 | 0.1 | --- |
| Lag Period, Stage 3 | $-5\times10^{-3}$ | $(-10,3)\times10^{-3}$ | 2 | 0.2 | --- |
| Generation Time, Stage 1 | 0.08 | (0.07,0.08) | 330 | <0.0001 | 6 |
| Generation Time, Stage 2 | 0.01 | $(2,23)\times10^{-3}$ | 0.9 | 0.1 | --- |
| Generation Time, Stage 3 | 0.09 | (0.08,0.10) | 375 | <0.0001 | 5 |

(a) CI = Confidence Interval for the coefficient

#### 6.3.1.3 One-Dimensional Simulation of Variability and Uncertainty

This section presents the results of regression analysis applied to a one-dimensional probabilistic simulation in which variability and uncertainty are co-mingled, based upon the case study scenario described in Section 3.3.3.

Table 6-11 summarizes the results of application of regression analysis to the growth estimation part for the co-mingled simulation of variability and uncertainty. The inputs are ranked based on the magnitude of the regression coefficients. Rankings are presented for statistically significant inputs with Pr>F less than 0.05. The F values in Table 6-11 indicate that there are no statistically significant effects for the storage time, the storage temperature, the generation time at stage 2, and lag period at stage 3.

Based upon the magnitude of the coefficients for the statistically significant inputs, the storage temperature at stage 3, the storage time at stage 3, the storage temperature at stage 1, and the storage time at stage 1 are the top four inputs. In order to evaluate the robustness of the estimated rankings, the 95 percent confidence intervals are estimated for each coefficient.

Estimated confidence intervals for regression coefficients indicate that the rankings for the top four inputs are unambiguous. For example, the confidence intervals for the storage temperature and time at stage 3 do not overlap. The confidence intervals for fifth and sixth inputs have overlap indicating that they are of comparable importance. However, both of these inputs have coefficients that are substantially smaller than inputs with higher ranks. Therefore, these two inputs are of little importance compared to the other ranked inputs. The maximum density and the lag period at stage 1 have the same magnitude of coefficients and confidence intervals. Hence, both of these inputs were ranked seventh.

If F values were used instead of regression coefficients as a basis for ranking, the results would be similar but not identical. The largest F value is associated with the input that has the second largest regression coefficient. In addition, input that has rank third based on the magnitude of its coefficient has rank fourth based on its F value. Thus, the rankings of the top two inputs and third and fourth inputs would be in reverse order. With this exception, all of the other rankings would remain approximately the same.

#### 6.3.1.4 Summary and Comparison of the Results of Regression Analysis in the Growth Estimation Part

In Sections 6.3.1.1 to 6.3.1.3 regression analysis was applied to three datasets considering variability only, variability for different uncertainty realizations, and co-mingled variability and uncertainty in inputs. In this section the rankings based on these analyses are summarized and compared. Table 6-12 gives the ranks for each input based on analyses in Sections 6.3.1.1 to 6.3.1.3.

The key similarities among the three probabilistic simulations were with respect to identification of the most important input, a group of three inputs with secondary importance, and a group of three inputs with moderate importance. There are some differences in rankings for the least importance inputs based upon these three simulations. The first and third simulations presented a complete agreement regarding ranking of inputs. All three probabilistic analysis methods resulted in identification of storage temperature at stage 3 as clearly the most important input. Storage time at stage 3, storage temperature at stage 1, and storage time at stage 1 were identified in the group of secondary importance inputs by all three simulations. All three simulations selected generation time at stages 3 and 1 and lag period at stage 1 in the group of

Table 6-12. Summary of the Regression Analysis Results Based on Variability Only, Variability for Different Uncertainty Realizations, and Co-mingled Variability and Uncertainty Analyses

| Variable | Ranks | | |
|---|---|---|---|
| | Analysis 1[(1)] | Analysis 2[(2)] | Analysis 3[(3)] |
| Storage Temperature, Stage 1 | 3 | 3.1 | 3 |
| Storage Temperature, Stage 2 | --- | 8.5 | --- |
| Storage Temperature, Stage 3 | 1 | 1.4 | 1 |
| Storage Time, Stage 1 | 4 | 3.5 | 4 |
| Storage Time, Stage 2 | --- | 10.8 | --- |
| Storage Time, Stage 3 | 2 | 2.7 | 2 |
| Maximum Density | 7 | 10.1 | 7 |
| Lag Period, Stage 1 | 7 | 8.9 | 7 |
| Lag Period, Stage 2 | --- | 9.7 | --- |
| Lag Period, Stage 3 | --- | 9.7 | --- |
| Generation Time, Stage 1 | 6 | 7.2 | 6 |
| Generation Time, Stage 2 | --- | 9.0 | --- |
| Generation Time, Stage 3 | 5 | 6.4 | 5 |

(1) Ranks based on the variability only analysis.
(2) Mean ranks based on the variability for different uncertainty realizations analysis.
(3) Ranks based on the one-dimensional co-mingled variability and uncertainty analysis.

moderate importance inputs. Inputs related to stage 2 were identified as statistically insignificant by first and third simulations. The second simulation considered mean ranks between 9.0 and 10.8 for these inputs.

There were some inputs for which the rankings appear to be different based upon the three simulation methods. For example, the storage temperature at stage 2 that has a mean rank of 8.5 in the second simulation, while it was identified as statistically insignificant by other two methods. Moreover, maximum density was identified as the seventh important input by first and third simulations, while it has mean rank of 10.1 based on the second simulation.

Although some of the inputs in least importance group had different rankings when comparing the three simulation methods, the most important finding is that the ranking of the top input, secondary importance inputs, and moderate importance inputs were essentially the same for all three approaches. This implies that any of the three approaches could be used alone in order to make a distinction between the top, secondary and moderate inputs.

The generation times at stages 1 and 3 are identified in the group of minor importance inputs by all three simulations, while the lag period at stage 1 and the maximum density are

classified into the group of minor importance inputs just by the variability only and one-dimensional co-mingled variability and uncertainty simulations.

Moreover, the variability only and one-dimensional co-mingled variability and uncertainty analysis have agreement on the storage temperature, the storage time, lag period and generation time at stage 2 and lag period at stage 3 as inputs with no statistically significant effects.

### 6.3.2 Correlation Coefficients in the Growth Estimation Part

The objective of this section is to present the results of applying correlation coefficients, as a method for the sensitivity analysis, to the growth estimation part. The growth estimation part was selected for this purpose because the two-dimensional probabilistic simulation, as explained in Section 3.2.3, makes the application of correlation coefficient challenging and manifests the capabilities of this method in identifying sensitive inputs in two-dimensional simulations. The details of the methodology for correlation coefficients methods are provided in Section 2.2.4. Two methods for correlation coefficient analysis including, Pearson (sample) and Spearman (rank) techniques are considered. Sections 6.3.2.1 and 6.3.2.2 present the results for these two methods, respectively.

#### 6.3.2.1 Pearson Correlation Coefficient in the Growth Estimation Part

The application of Pearson correlation coefficients to the two-dimensional probabilistic simulation involves sensitivity analysis for each of 100 uncertainty iterations. Within each uncertainty iteration, 650 samples were generated to represent variability for each input. Thus, correlation coefficients were generated 100 times.

The results of the 100 analyses with Pearson correlation coefficients are summarized in Table 6-13. The table includes the mean correlation coefficients and the 95 percent probability range of coefficients over the 100 simulations. The percentage of the 100 simulations that produced a statistically significant coefficient is quantified. Furthermore, the mean rank and the range of ranks are given for each input. The inputs included in the Pearson correlation coefficient analysis were the same as the variables listed in Table 6-9.

The mean ranks indicate that the storage temperature at stage 3 is the most important input. There is 100 percent probability that this input is identified as statistically significant in the uncertainty realizations. The mean ranks for the storage time at stage 3 and the generation time at stage 3 are estimated as 3.9 and 3.8, respectively, indicating that on average the output has

Table 6-13. Summary of the Pearson Correlation Coefficient Results for Two-Dimensional Variability Simulation for Different Uncertainty Realizations

| Variable | Mean Correlation Coefficient | 95% Probability Range of Coefficients | Frequency | Mean Rank | Range of Rank |
|---|---|---|---|---|---|
| Temp1 | 0.269 | (0.030,0.525) | 93 | 4.3 | 1-13 |
| Temp2 | 0.008 | (-0.070,0.102) | 7 | 10.9 | 5-13 |
| Temp3 | 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 |
| MD | 0.027 | (-0.052,0.095) | 10 | 10.6 | 6-13 |
| LP1 | -0.169 | (-0.317,-0.022) | 83 | 7.0 | 3-13 |
| LP2 | -0.009 | (-0.104,0.071) | 7 | 10.8 | 5-13 |
| LP3 | -0.311 | (-0.45,-0.146) | 100 | 4.8 | 2-8 |
| GT1 | -0.168 | (-0.328,0.00) | 79 | 7.1 | 2-13 |
| GT2 | -0.006 | (-0.092,0.065) | 7 | 10.7 | 6-13 |
| GT3 | -0.339 | (-0.496,-0.162) | 100 | 3.8 | 2-8 |

approximately similar sensitivity to these inputs. For these inputs the probability of being statistically significant is 98 percent or more. However, although these inputs have approximately similar average rankings indicating that they are of comparable importance to each other, they are less important than the storage temperature at stage 3. The storage temperature and time at stage 1, and lag period at stage 3 have mean ranks of 4.3, 4.7, and 4.8, respectively. These inputs are considered to be of secondary importance. The lag period at stage 1 and the generation times at stage 1 have mean ranks between 7.0 and 7.1 with probability of being statistically significant of 83 and 79 percent, respectively, indicating that the output has similar sensitivity to these inputs. The output has the lowest sensitivity to inputs corresponding to the second stage, transportation, and the maximum density. The mean ranks for these inputs vary between 10.6 and 10.9. These inputs had statistically significant effects in only 7 to 10 percent of the uncertainty realizations.

In order to visualize the results of the sensitivity analysis, the complementary cumulative distribution function (CCDF) of the rank is given for each input based upon the 100 uncertainty realizations in Figures 6-11 to 6-13. Figure 6-10 displays the CCDFs for six inputs that have the highest average ranks among all of the inputs included in the analysis. These inputs include

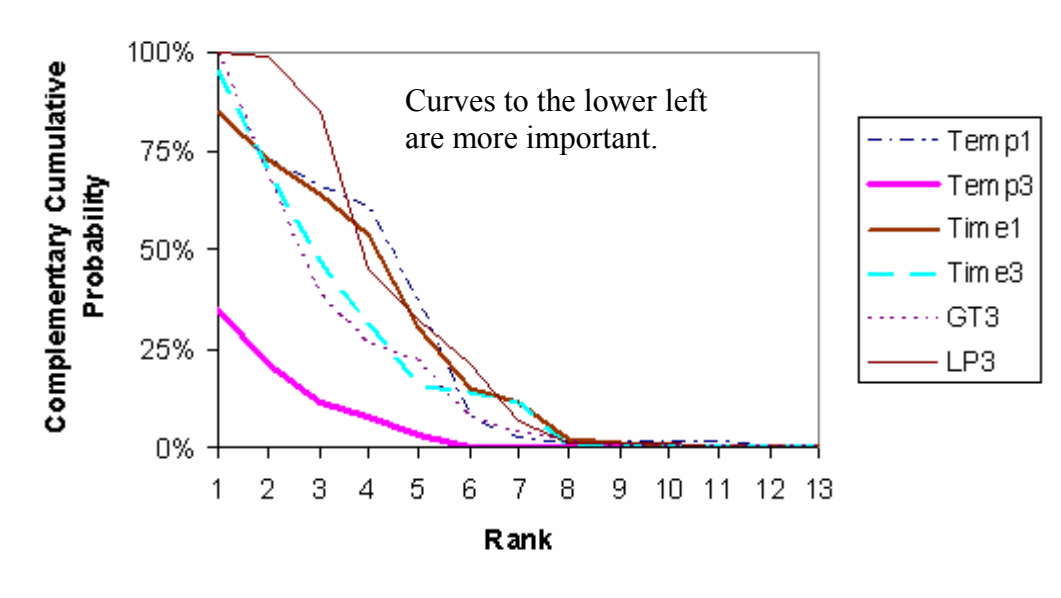



Figure 6-11. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected Inputs Based Upon Pearson Correlation: Storage Temperature at Stages 1and 3 ($Temp_1$ and $Temp_3$); Storage Time at Stages 1 and 3 ($Time_1$ and $Time_3$); and Generation Time at Stage 3 (GT3).

storage time and temperature at stages 1 and 3, and lag period and generation time at stage 3. The CCDF for the storage temperature at stage 3 indicates that for 35 percent of the simulations, the rank was worse than one, which implies that the rank was equal to one for 65 percent of the simulations. Furthermore, the storage temperature at stage 3 was ranked sixth or higher for 100 percent of the simulations. In contrast, storage time at stage 1 was ranked first for 15 percent of the simulations and was ranked sixth or higher for 80 percent of the simulations. The frequencies of being the most important input for storage temperature at stage 1 and storage time at stage 3 are 15 and 5 percent, respectively. Thus, although the storage temperature at stage 3 has the highest frequency of a rank of one, there is some ambiguity regarding which of the other three inputs is the second most important.

Figure 6-12 shows the CCDFs for two inputs that have average ranks between 7.0 and 7.1, while Figure 6-13 depicts the CCDFs for five inputs of approximately minor importance with average ranks varying between 10.6 and 10.9. These latter inputs were mostly identified as not statistically significant.

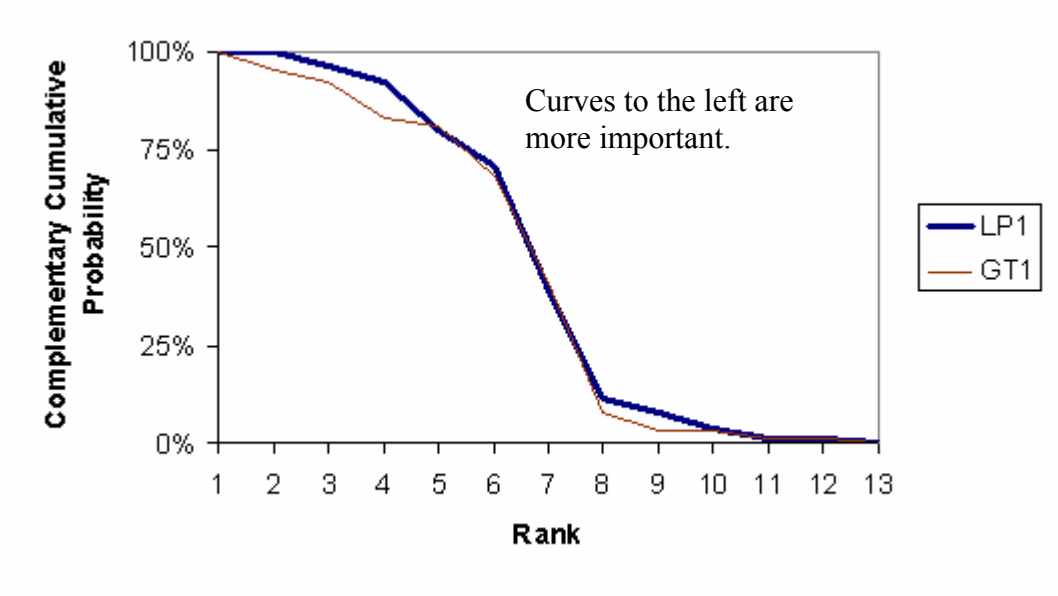



Figure 6-12. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected Inputs Based Upon Pearson Correlation: Lag Period at Stages 1 and 3 (LP1 and LP3); and Generation Time at Stage 1 (GT1).

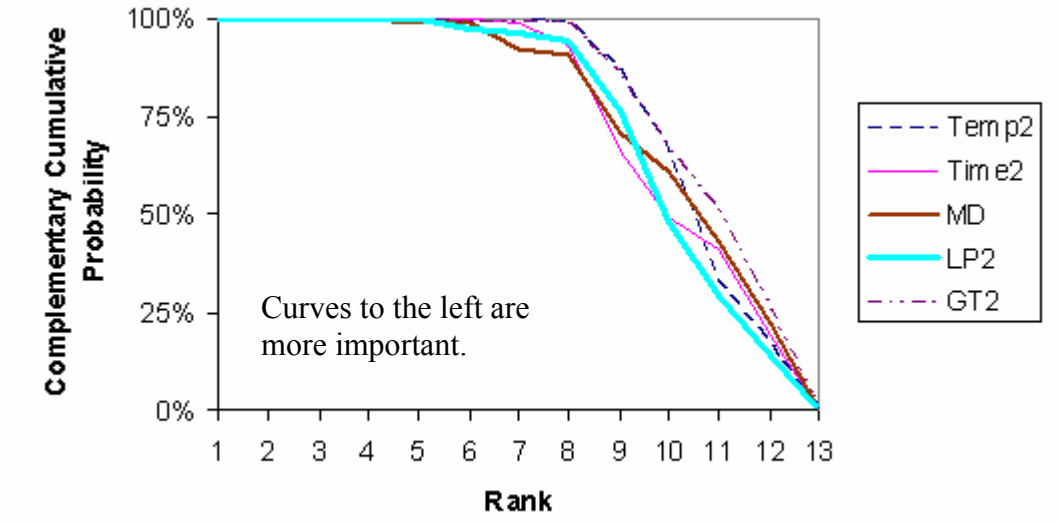



Figure 6-13. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected Inputs Based Upon Pearson Correlation: Storage Time and Temperature, Lag Period, and Generation Time at Stage 2 (Time2, Temp2, LP2, and GT2); and Maximum Density (MD).

The results shown in the three figures indicate that there are approximately four groups of inputs. These groups include: (1) the most important input of storage temperature at stage 3; (2) inputs of secondary importance, including storage time at stages 1 and 3 and storage temperature at stage 1, and generation time and lag period at stage 3; (3) inputs of tertiary importance, including generation time and lag period at stage 1; and (4) inputs of minor or no importance, including inputs corresponding to the second stage (i.e., transportation) and the maximum density.

6.3.2.2 Spearman Correlation Coefficients in the Growth Estimation Part

This section presents the results based upon Spearman (rank) correlation coefficients applied to the growth estimation part using a two-dimensional probabilistic framework. Details regarding the Spearman correlation coefficient technique are provided in Section 2.2.4.

The results of the 100 analyses with Spearman correlation coefficients are summarized in Table 6-14 similar to the summary in Table 6-13 for the Pearson correlation coefficients. Inputs to the analysis are same as the variables listed in Table 6-9.

The mean ranks indicate that the storage time at stage 3 is the most important input. There is 100 percent probability that this input is identified as statistically significant in the uncertainty realizations. The storage time at stage 1 is also identified a top input with a mean rank and a range of ranks approximately equal to the most important input. Hence, the output may have comparable sensitivity to the storage time at stages 1 and 3.

There are four inputs with mean ranks varying between 3.3 and 5.4. These inputs include lag period, generation time, storage temperature at stage 3 and lag period at stage 1. These inputs are categorized to be of secondary importance and are statistically significant. Two inputs, of tertiary importance, are storage temperature and generation time at stage 1 with mean ranks of 6.6 and 7.2, respectively. There are five inputs for which most simulations were not statistically significant and the average ranks were low, including storage temperature, storage time, lag period, and generation time at stage 1 and the maximum density. These inputs were deemed to be unimportant.

In order to visualize the results of the sensitivity analysis, the complementary cumulative distribution function (CCDF) of the rank is given for each input based upon the 100 uncertainty realizations. Figure 6-14 shows the CCDFs for the six inputs that have the highest average ranks. Storage times at stages 1 and 3 have a comparable probability of being ranked first. The other

Table 6-14. Summary of the Spearman Correlation Coefficient Results for Two-Dimensional Variability Simulation for Different Uncertainty Realizations

| Variable | Mean Corr. Coefficient | 95% Probability Range of Coefficients | Frequency | Mean Rank | Range of Rank |
|---|---|---|---|---|---|
| Temp1 | 0.178 | (0.070,0.279) | 97 | 6.6 | 1-9 |
| Temp2 | 0.024 | (-0.050,0.110) | 11 | 10.8 | 6-13 |
| Temp3 | 0.263 | (0.132,0.398) | 100 | 5.1 | 1-8 |
| Time1 | 0.443 | (0.258,0.619) | 100 | 1.9 | 1-5 |
| Time2 | 0.013 | (-0.059,0.081) | 7 | 11.1 | 7-13 |
| Time3 | 0.436 | (0.232,0.585) | 100 | 1.7 | 1-6 |
| MD | 0.007 | (-0.072,0.085) | 6 | 11.1 | 7-13 |
| LP1 | -0.234 | (-0.335,-0.101) | 99 | 4.9 | 2-10 |
| LP2 | -0.030 | (-0.114,0.049) | 14 | 10.7 | 6-13 |
| LP3 | -0.310 | (-0.418,-0.188) | 100 | 3.3 | 2-7 |
| GT1 | -0.166 | (-0.280,-0.043) | 90 | 7.2 | 3-13 |
| GT2 | -0.024 | (-0.116,0.053) | 8 | 10.9 | 8-13 |
| GT3 | -0.255 | (-0.381,-0.140) | 100 | 5.4 | 3-8 |

four inputs are never identified as the most important input in the analysis. The probability that the storage time at stage 1 has a rank worse than 5 is zero. In contrast several of the other inputs shown in the figure have ranks as low as eight.

In order to visualize the results of the sensitivity analysis, the complementary cumulative distribution function (CCDF) of the rank is given for each input based upon the 100 uncertainty realizations. Figure 6-14 shows the CCDFs for the six inputs that have the highest average ranks. Storage times at stages 1 and 3 have a comparable probability of being ranked first. The other four inputs are never identified as the most important input in the analysis. The probability that the storage time at stage 1 has a rank worse than 5 is zero. In contrast several of the other inputs shown in the figure have ranks as low as eight.

Figure 6-15 depicts a group of inputs with tertiary importance, including generation time and storage temperature at stage 1. There is approximately 84 percent probability that these inputs have ranks worse than five. Figure 6-16 depicts a set of five inputs that are comparatively insensitive. For these inputs, there is less than 2 percent probability of having a rank higher than 8. These inputs mostly correspond to the second stage (i.e., transportation) indicating that the transportation stage does not have a significant effect on the output in the growth estimation part.

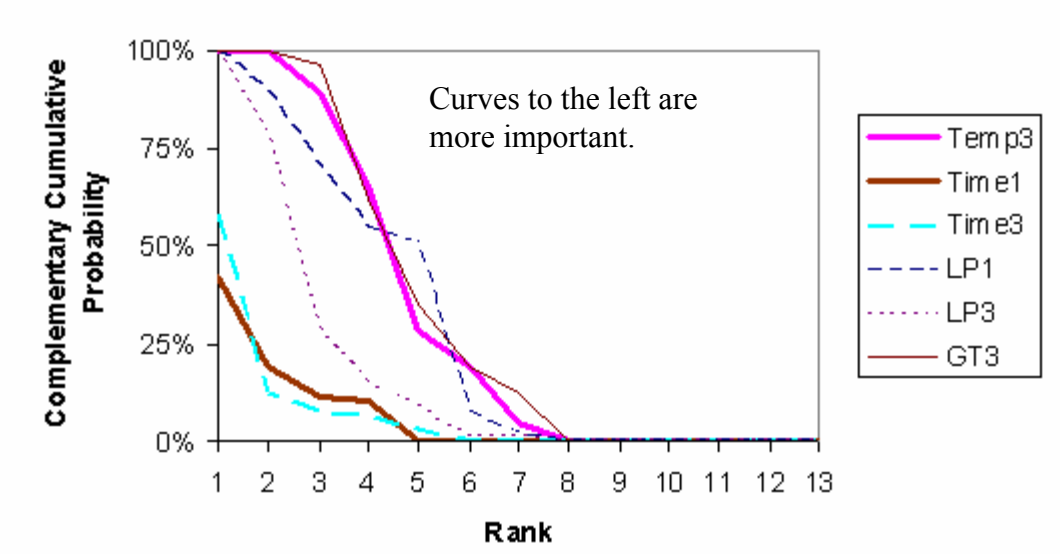



Figure 6-14. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected Inputs Based Upon Spearman Correlation: Storage Time, Temperature, Lag Period, and Generation Time at Stage 3 (Time3, Temp3, LP3, and GT3); Storage Time and Lag Period at Stage 1 (Time1 and LP1).

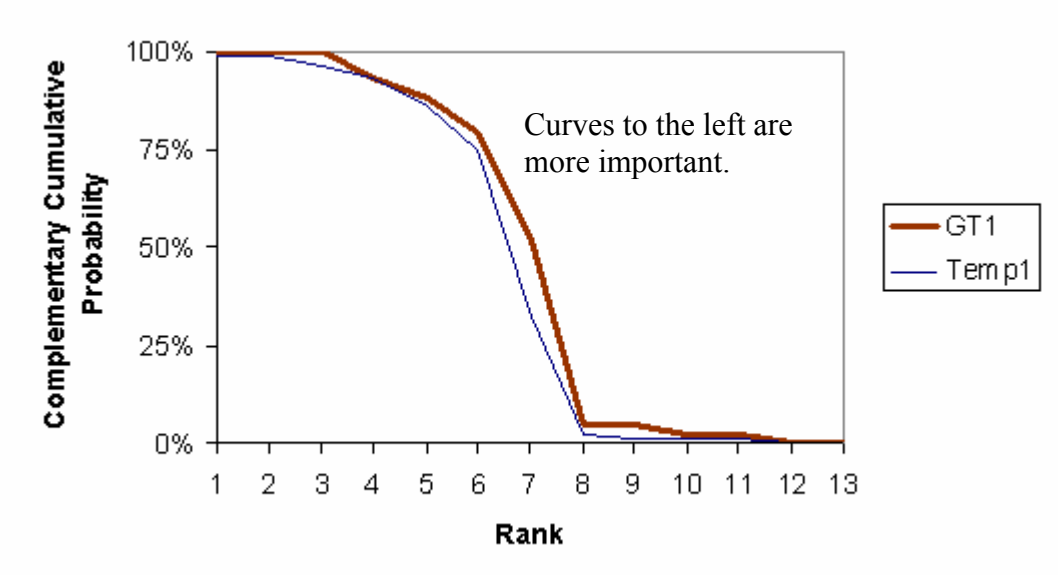



Figure 6-15. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected Inputs Based Upon Spearman Correlation: Generation Time at Stage 1 (GT1); and Storage Temperature at Stage 1 (Temp1).

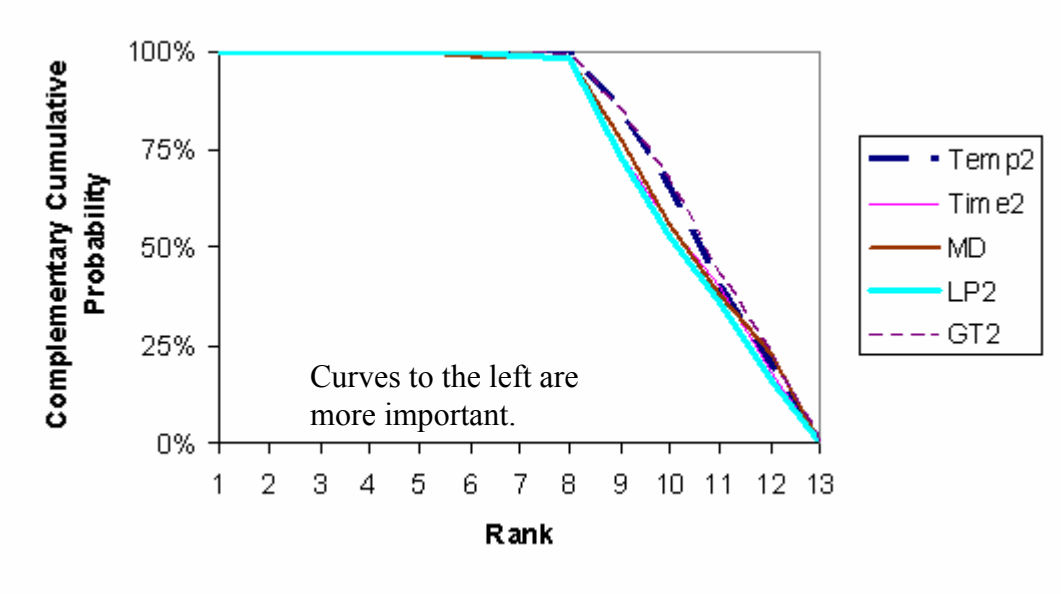



Figure 6-16. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected Inputs Based Upon Spearman Correlation: Storage Temperature, Storage Time, Lag Period, and Generation Time at Stage 2 (Temp2, Time2, LP2, and GT2); and Maximum Density (MD).

#### 6.3.2.3 Comparison of the Results from the Pearson and Spearman Correlation Coefficients

In Sections 6.3.2.2 and 6.3.2.3 Pearson and Spearman correlation coefficient methods were applied to the two dimensional framework of variability and uncertainty in the growth estimation part. This section summarizes and compares the rankings based on these analyses. Table 6-15 gives the mean ranks for each input based on the two analyses.

The key similarity among Pearson and Spearman correlation coefficient methods was with respect to the identification of the least important inputs. Both methods resulted in identification of second stage inputs as the least important. In addition, maximum density was also identified in a group of least important inputs by both methods. However, the two differ in the identification of the most important and the group of secondary importance inputs. For instance, although storage times at stages 1 and 3 were selected as the most important inputs based upon Spearman method, these inputs did not achieve a mean rank greater than 3.7 with the Pearson method. The results from the Spearman-based approach are considered to be more

Table 6-15. Summary of the Pearson and Spearman Correlation Coefficient Analysis Based Upon Variability Analysis for 100 Uncertainty Realizations in Growth Estimation Part

| Variable | Ranks | |
|---|---|---|
| | Pearson | Spearman |
| Storage Temperature, Stage 1 | 4.3 | 6.6 |
| Storage Temperature, Stage 2 | 10.9 | 10.8 |
| Storage Temperature, Stage 3 | 1.6 | 5.1 |
| Storage Time, Stage 1 | 4.7 | 1.9 |
| Storage Time, Stage 2 | 10.9 | 11.1 |
| Storage Time, Stage 3 | 3.9 | 1.7 |
| Maximum Density | 10.6 | 11.1 |
| Lag Period, Stage 1 | 7.0 | 4.9 |
| Lag Period, Stage 2 | 10.8 | 10.7 |
| Lag Period, Stage 3 | 4.8 | 3.3 |
| Generation Time, Stage 1 | 7.1 | 7.2 |
| Generation Time, Stage 2 | 10.7 | 10.9 |
| Generation Time, Stage 3 | 3.8 | 5.4 |

accurate in this case because the growth process is not linear. Although the equation used for the growth process typically provides a monotonic association between the estimated growth and the inputs, there are some conditions considered in the model in which growth does not change if particular inputs are below a threshold or if a maximum growth rate is achieved. In particular, two conditions in the model include a comparison of the estimated growth with maximum density and a comparison of the storage time with the available lag period at each stage. These conditions enforce a constant response in the growth estimation part. For example, if the storage time is less than the available lag period, no growth is estimated for the number of *E. coli* organisms in the ground beef servings. Hence, in these cases, an increase in the storage time is not accompanied by a simultaneous monotonic change in the estimated growth. This condition to the model prevents the model from responding in a completely monotonic pattern. Therefore, the assumption of monotonic association between the output and the input considered in the Spearman correlation coefficient method is not completely valid in this case.

### 6.3.3 Rank Regression in the Growth Estimation Part

This section presents the results of the rank regression on the growth estimation part of the preparation module. Application of the standardized regression analysis to the growth estimation part in Section 6.3.1 was accompanied with a moderate $R^2$ value of approximately 0.5. This value of $R^2$ implies that underlying assumption of linear relationship between the

Table 6-16. Summary of the Rank Regression Results for Two-Dimensional Variability Simulation for Different Uncertainty Realizations (Mean R2 = 0.55)

| Variable | Mean Coefficient | 95% Probability Range of Coefficients | Frequency | Mean Rank | Range of Rank |
|---|---|---|---|---|---|
| Temp1 | 0.016 | (-0.053,0.085) | 7 | 9.3 | 4-13 |
| Temp2 | 0.012 | (-0.102,0.120) | 9 | 8.1 | 4-13 |
| Temp3 | 0.045 | (-0.026,0.102) | 18 | 8.0 | 4-13 |
| Time1 | 0.401 | (0.225,0.555) | 100 | 1.5 | 1-3 |
| Time2 | 0.017 | (-0.025,0.062) | 6 | 10.4 | 5-13 |
| Time3 | 0.381 | (0.190,0.539) | 100 | 1.6 | 1-3 |
| MD | 0.001 | (-0.037,0.041) | 3 | 11.2 | 7-13 |
| LP1 | -0.175 | (-0.268,-0.078) | 100 | 3.6 | 2-7 |
| LP2 | -0.024 | (-0.105,0.050) | 13 | 9.0 | 5-13 |
| LP3 | -0.199 | (-0.300,-0.097) | 100 | 3.4 | 2-8 |
| GT1 | -0.027 | (-0.097,0.051) | 10 | 9.0 | 5-13 |
| GT2 | 0.008 | (-0.086,0.116) | 9 | 8.6 | 4-13 |
| GT3 | 0.008 | (-0.086,0.116) | 43 | 7.0 | 4-13 |

output and inputs to this part of the model is not strongly valid. Hence, rankings of sensitivity based upon the magnitude of the standardized regression coefficients may be unreliable. Rank regression can more adequately address nonlinear monotonic relationships than sample regression. Details regarding rank regression method are provided in Section 2.2.5.

Rank regression was applied to the two-dimensional simulation of variability for 100 uncertainty realizations in the growth estimation part. The characteristics of the simulation regarding the number of variability and uncertainty iterations are the same as those in Sections 6.3.1, 6.3.2.1, and 6.3.2.2. The results of the 100 analyses with the rank regression technique are summarized in Table 6-16, similar to earlier summaries for other methods. The inputs included in the rank regression method for the two-dimensional simulation were the same as those listed in Table 6-9.

The mean ranks indicate that both the storage times at stages 1 and 3 are the most important inputs. Both inputs have range of ranks between 1 and 3 and were statistically significant in all the uncertainty realizations. The lag periods at stages 1 and 3 are considered to be of secondary importance inputs with mean ranks ranges between 3.42 and 3.62. A group of five inputs were deemed to be of minor importance, including generation times at stages 1, 2, and 3, and storage temperatures at stages 2 and 3. These inputs were statistically significant in

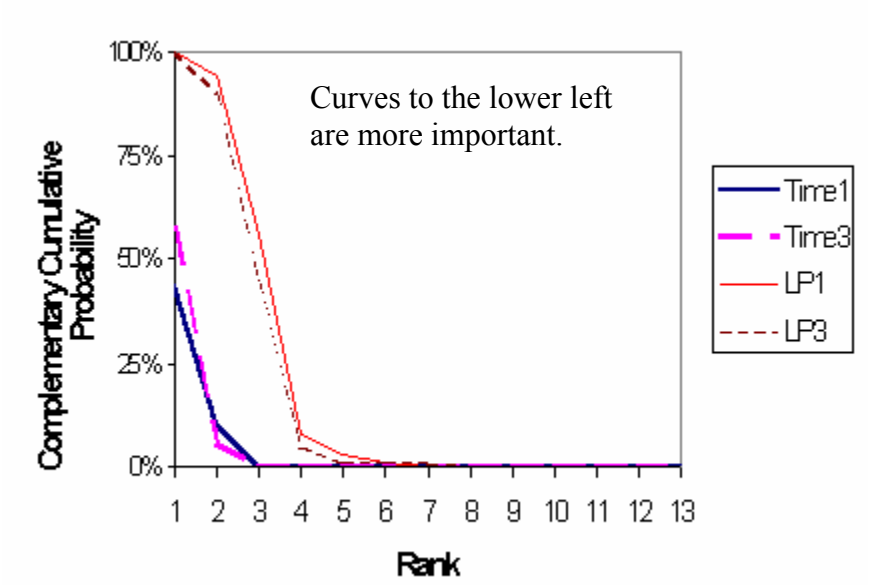



Figure 6-17. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected Inputs Based Upon Rank Regression: Storage Time at Stages 1 and 3 (Time1 and Time3); and Lag Period at Stages 1 and 3 (LP1 and LP3).

only 9 to 43 percent of the uncertainty realizations. Finally, four inputs were grouped as unimportant, including storage temperature at stage 1, storage time and lag period at stage 2, and maximum density. These inputs had low average ranks and typically were statistically insignificant.

The complementary cumulative distribution function (CCDF) of the rank is given for each input. Figure 6-17 displays the CCDFs for four inputs that have the highest average ranks. The CCDF for the storage time at stage 3 indicates that the rank was equal to one for 43 percent of the simulations and was ranked three or higher for 100 percent of the simulations. Storage time at stage 1 was ranked first in 57 percent of uncertainty realizations and was always higher than third. The other two inputs shown were never identified as the most important inputs and typically had much lower ranks.

Figure 6-18 shows the CCDFs for five inputs that have lower importance than those depicted in Figure 6-17. There was 100 percent probability that these inputs had ranks not better than third, and a probability of 23 to 52 percent of having a rank worse than 8.

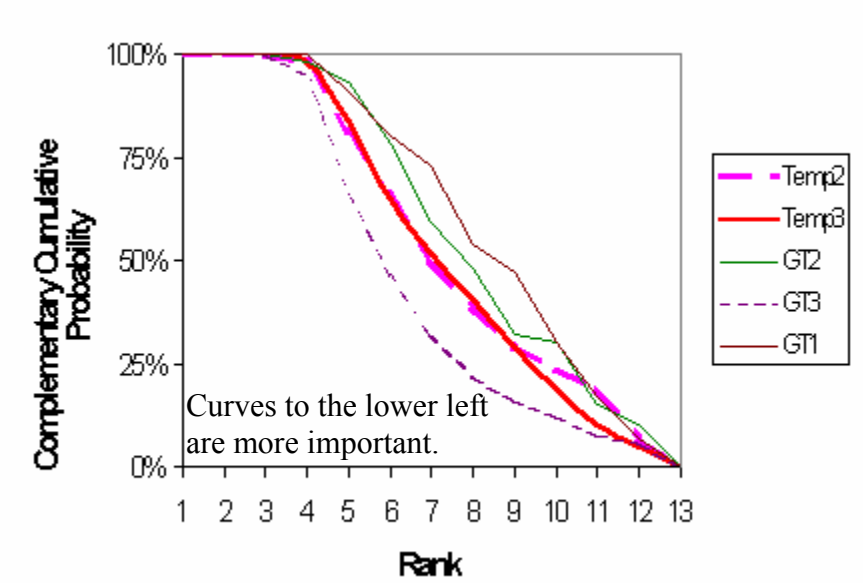



Figure 6-18. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected Inputs Based Upon Rank Regression: Storage Temperature at Stages 2 and 3 (Temp2 and Temp3); and Generation Time at Stages 1, 2, and 3 (GT1, GT2, and GT3).

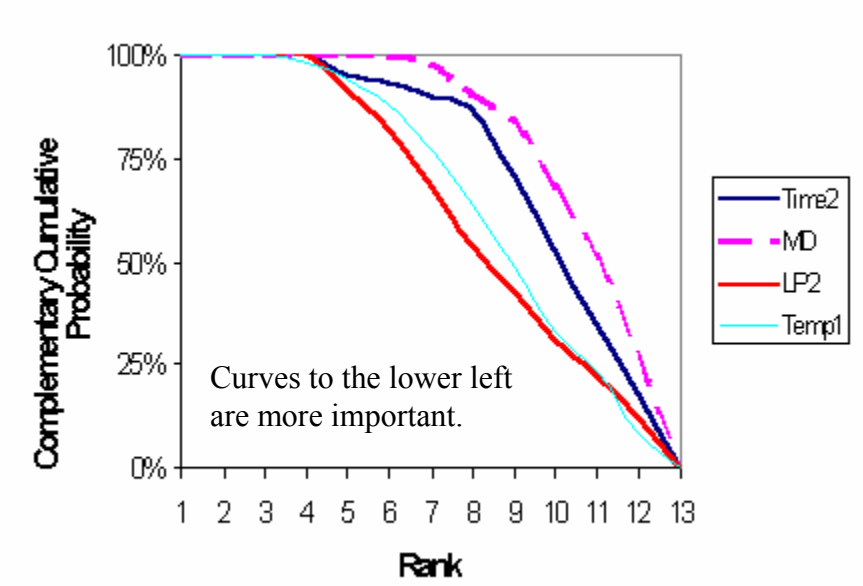



Figure 6-19. Complementary Cumulative Distribution Functions (CCDFs) of Uncertainty in the Rank of Selected Inputs Based Upon Rank Regression: Storage Time at Stages 2 (Time 2); Maximum Density (MD); Lag Period at Stage 2; and Storage Temperature at Stage 1 (Temp1).

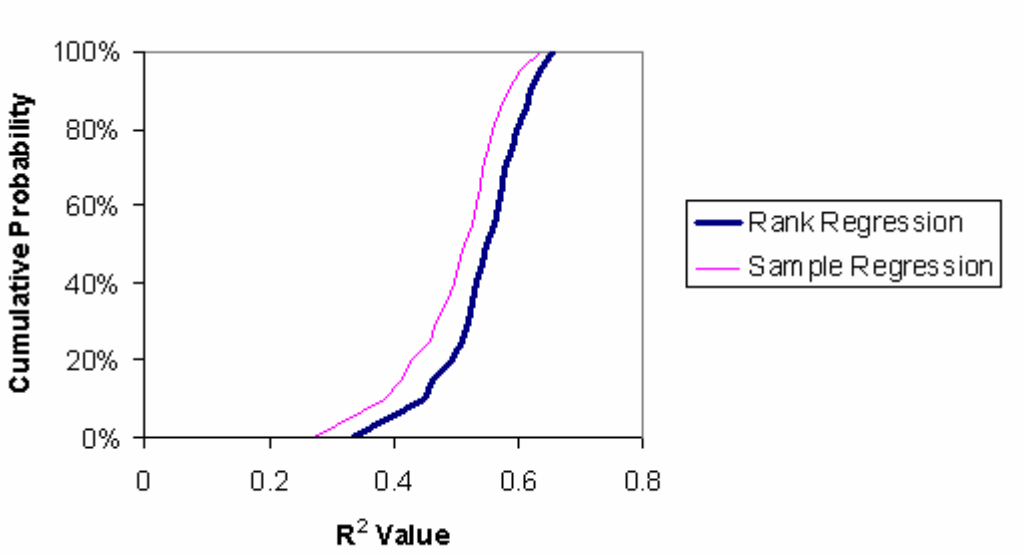



Figure 6-20. $R^2$ Distribution for Two-Dimensional Rank Regression in the Growth Estimation Part.

Figure 6-19 depicts the four inputs with the lowest sensitivity. There was 100 percent probability that the ranks for these inputs were not better than fourth. Moreover, the probability that these inputs had ranks worse than 10 ranges between 27 and 69 percent.

Figure 6-20 depicts the cumulative probability function (CDF) for the 100 $R^2$ values obtained in the two-dimensional simulation. The $R^2$ value for the rank regression method varied between 0.31 and 0.67 with an average of 0.55. This average is not substantially better that that obtained using standardized sample linear regression, which had an average $R^2$ value of 0.50. Although rank regression is robust to the underlying assumption of linearity, this method assumes that there is a monotonic relationship between the output and inputs to the model. Section 6.3.2.3 presents a discussion regarding the reason that in the growth estimation part the model has only a partial monotonic response with respect to variation of the inputs. In particular, for some ranges of values of specific inputs, either zero growth or maximum growth is estimated. Therefore, there is not a substantial improvement in the $R^2$ value using the rank regression method.

### 6.3.4 Summary and Comparison of Results in the Growth Estimation Part

In Sections 6.3.1 to 6.3.3 standardized linear regression, correlation coefficients, and rank regression were applied to the growth estimation part. This section provides a summary and comparison of the results based on these methods as shown in Table 6-17.

Table 6-17. Summary of the Mean Ranks Based on Standardized Linear Regression, Correlation Coefficients, and Rank Regression in the Growth Estimation Part

| Variable | Regression Analysis | Correlation Coefficients | | Rank Regression |
|---|---|---|---|---|
| | | Pearson | Spearman | |
| Storage Temperature, Stage 1 | 3.1 | 4.3 | 6.6 | 9.3 |
| Storage Temperature, Stage 2 | 8.5 | 10.9 | 10.8 | 8.1 |
| Storage Temperature, Stage 3 | 1.4 | 1.6 | 5.1 | 8.0 |
| Storage Time, Stage 1 | 3.5 | 4.7 | 1.9 | 1.5 |
| Storage Time, Stage 2 | 10.8 | 10.9 | 11.1 | 10.4 |
| Storage Time, Stage 3 | 2.7 | 3.9 | 1.7 | 1.6 |
| Maximum Density | 10.1 | 10.6 | 11.1 | 11.2 |
| Lag Period, Stage 1 | 8.9 | 7.0 | 4.9 | 3.6 |
| Lag Period, Stage 2 | 9.7 | 10.8 | 10.7 | 9.0 |
| Lag Period, Stage 3 | 9.7 | 4.8 | 3.3 | 3.4 |
| Generation Time, Stage 1 | 7.2 | 7.1 | 7.2 | 9.0 |
| Generation Time, Stage 2 | 9.0 | 10.7 | 10.9 | 8.6 |
| Generation Time, Stage 3 | 6.4 | 3.8 | 5.4 | 7.0 |

According to the results provided in Table 6-17, the two sample-based methods of standardized linear regression analysis and Pearson correlation coefficients produced approximately similar ranking for inputs. There is also similarity in rankings between the two ranked-based methods of rank regression and Spearman correlation coefficients. Generally, results according to the rank-based techniques for sensitivity analysis are different from those of the method methods based on the sample data. The differences between sample and rank based techniques are more apparent with respect to the inputs to which model has higher sensitivity. For example, while storage temperature at stage 3 was identified as the most important input using the standardized regression analysis and sample (Pearson) correlation coefficients methods, this input was attributed low mean ranks of 5.1 and 8.0 using rank (Spearman) correlation coefficients and rank regression methods, respectively. All methods approximately identified the same inputs that have low or no importance. For example, inputs associated with stage 2 and maximum density were attributed low mean ranks between 8.1 and 11.2.

### 6.3.5 Regression Analysis for Variability in the Serving Contamination Part

The serving contamination part of the preparation module is explained in Section 3.4.3.1. Inputs include the ground beef consumption type, serving size, eating location, consumer age, and grinder contamination. Distributions for these inputs are summarized in Table 3-12. The output in this part is the mean serving contamination. For this part there is a one-dimensional

Table 6-18. The Regression Analysis Results for the Serving Contamination Part in Summer ($R^2$ = 0.11)

| Variable | Coefficient | 95% CI [a] | F Value | Pr>F | Rank |
|---|---|---|---|---|---|
| Ground Beef Consumption Type | --- | --- | 1 | 0.4 | --- |
| Eating Location | --- | --- | 36 | <0.0001 | --- |
| Consumer Age | --- | --- | 0.5 | 0.7 | --- |
| Serving Size | 0.12 | (0.11,0.13) | 970 | <0.0001 | 2 |
| Grinder Contamination | 0.32 | 0.32[b] | 7300 | <0.0001 | 1 |

a) CI = Confidence Interval for the coefficient
b) The interval for this coefficient is so tight that it appears as 0.32 to 0.32 when it is rounded to two decimal places.

Table 6-19. The Regression Analysis Results for the Serving Contamination Part in Winter ($R^2$ = 0.05)

| Variable | Coefficient | 95% CI [a] | F Value | Pr>F | Rank |
|---|---|---|---|---|---|
| Ground Beef Consumption Type | --- | --- | 4 | 0.03 | --- |
| Eating Location | --- | --- | 9 | 0.003 | --- |
| Consumer Age | --- | --- | 0.5 | 0.7 | --- |
| Serving Size | 0.05 | (0.04,0.05) | 130 | <0.0001 | 2 |
| Grinder Contamination | 0.13 | (0.12,0.14) | 1,200 | <0.0001 | 1 |

(a) CI = Confidence Interval for the coefficient

variability simulation with 65,000 iterations as explained in Section 3.3.3. The case scenario in the serving contamination part includes separate consideration of high and low prevalence seasons. The analyses are reported separately for these two seasons.

The results for the high prevalence season are given in Table 6-18. The rankings are based on the magnitude of the estimated regression coefficients for quantitative inputs. Rankings are presented for the statistically significant inputs with Pr>F less than 0.05. F values in Table 6-18 indicate that in the consumer age and the eating location are statistically significant. The grinder contamination is the most important quantitative input. The 95 percent confidence intervals for the quantitative inputs are estimated in order to evaluate the unambiguity of the rankings. The confidence intervals indicate that the rankings between two quantitative inputs are unambiguous because these intervals do not overlap.

If instead of the coefficient estimates, the magnitude of the F values is used as a criterion for ranking the inputs, the rankings would be the same. The $R^2$ for the linear regression model fitted to the dataset is 0.11. The estimated value of $R^2$ implies that the linear assumption for the

functional relation between the output and inputs explains only 11 percent of the variability in the output. Thus, ranking based on the magnitude of the linear regression coefficients may not be reliable. Section 11.1.3.3 presents the comparison of the results based on the standardized linear regression with other methods that do not assume specific functional relationships, such as ANOVA and CART. The results in Table 11-12 indicate that rankings based on the linear regression analysis are comparable to that of the CART method with respect to the selection of key inputs, while the rank order of first two inputs is reversed in ANOVA.

The results for the low prevalence season are given in table 6-19. The F values indicate that in winter the consumer age, ground beef consumption type, and the eating location are not statistically significant. Comparing the magnitude of the regression coefficients for quantitative inputs in Table 6-19 indicates that the grinder contamination is the most important input. The serving size is ranked as second input. The confidence intervals indicate that the rankings are unambiguous in winter because these intervals do not overlap.

Application of F values as a criterion for ranking inputs does not affect the previous ranking based on the magnitude of coefficient. The ranks based on the magnitude of the F values are robust, because F values differ substantially. For example, the F value for the grinder contamination differs from the F value for the serving size by a ratio of approximately 9.3.

For the low prevalence season results, the $R^2$ value of 0.05 is low, indicating that only a small amount of the variability in the output is addressed by the fitted linear model. However, as described in Section 11.1.3.3 and Table 11-13, rankings based on the linear regression analysis are comparable to that of the CART method with respect to the selection of key inputs, while the rank order of first two inputs is reversed in ANOVA.

Low values of $R^2$ in this part of the model indicate that standardized linear regression is not a reliable sensitivity analysis method in this case. Other variations of regression analysis might yield better results. Therefore, a comparison is made with other regression based approaches. In addition to the rank regression technique for capturing nonlinearity, application of higher order terms in the regression model and/or transformation techniques such as log scale transformation can improve the amount of variability of the output that can be captured by the fitted regression model. In order to illustrate this issue, rank regression, log scale transformation, and application of higher order terms in the regression model were implemented to the serving contamination part. $R^2$ values obtained from these analyses are used to compare these methods.

Table 6-20. Standardized Linear Regression Analysis Results for Serving Contamination in Hamburger Patties Consumed by People Between 25 and 64 Years Old in Summer ($R^2 = 0.10$)

| Variable | Coefficient | 95% CI [(a)] | F Value | Pr>F | Rank |
|---|---|---|---|---|---|
| Serving Size | 0.115 | (0.104,0.126) | 432 | <0.0001 | 2 |
| Grinder Contamination | 0.304 | (0.293,0.315) | 3,004 | <0.0001 | 1 |

(a) CI = Confidence Interval for the coefficient

Table 6-21. Rank Regression Analysis Results for Serving Contamination in Hamburger Patties Consumed by People Between 25 and 64 Years Old in Summer ($R^2 = 0.97$)

| Variable | Coefficient | 95% CI [(a)] | F Value | Pr>F | Rank |
|---|---|---|---|---|---|
| Serving Size | 0.228 | (0.225,0.232) | 18,189 | <0.0001 | 2 |
| Grinder Contamination | 0.959 | (0.956,0.962) | 321,151 | <0.0001 | 1 |

(a) CI = Confidence Interval for the coefficient

Table 6-22. Results of Standardized Regression Analysis with Log Transformation for Serving Contamination in Hamburger Patties Consumed by People Between 25 and 64 Years Old in Summer ($R^2 = 0.99$)

| Variable | Coefficient | 95% CI [(a)] | F Value | Pr>F | Rank |
|---|---|---|---|---|---|
| Serving Size | 0.197 | (0.196,0.199) | 94,500 | <0.0001 | 2 |
| Grinder Contamination | 0.976 | (0.975,0.978) | 2,313,600 | <0.0001 | 1 |

(a) CI = Confidence Interval for the coefficient

For these additional analyses, serving contamination in hamburger patties consumed by people between 25 and 64 years old in summer was selected as the output of interest. Hence, there were two inputs to the model including serving size and grinder contamination.

Table 6-20 summarizes the results of the linear regression. The $R^2$ value estimated for linear regression was 0.1 indicating that the linear assumption was likely to be inapplicable. Table 6-21 presents the results of the rank regression analysis. Rank regression improved the $R^2$ value to a high value of 0.97 indicating that fitting a monotonic model to the rank-ordered dataset captured almost all of the variability of the output. Table 6-22 summarizes the results of the linear regression when log scale transformation was performed. Log scale transformation produced an $R^2$ of 0.99. Hence, the log transformed model can capture almost all of the variation in the output.

Table 6-23 summarized the results of the regression analysis when higher order terms such as interaction, quadratic, and cubic terms were used in the fitted regression model. The $R^2$ value of 0.64 for this case is not as high as those of rank regression and regression with log scale

Table 6-23. Results of Regression Analysis with Higher Order Terms for Serving Contamination in Hamburger Patties Consumed by People Between 25 and 64 Years Old in Summer ($R^2 = 0.64$)

| Variable | Coefficient | 95% CI [(a)] | F Value | Pr>F | Rank |
|---|---|---|---|---|---|
| Serving Size (S) | $6 \times 10^{-3}$ | $(5.8,6.2) \times 10^{-3}$ | 375 | <0.0001 | 2 |
| Grinder Contamination (GC) | 1.156 | (1.133,1.179) | 9,692 | <0.0001 | 1 |
| S × GC | $1 \times 10^{-4}$ | $(0.9,1.1) \times 10^{-4}$ | 872 | <0.0001 | --- |
| S × S | $3 \times 10^{-7}$ | $(0.6,6.1) \times 10^{-7}$ | 5.6 | 0.02 | --- |
| S × S × S | 0 | --- | 0.2 | 0.5 | --- |
| GC × GC | 0.215 | (0.211,0.220) | 8,379 | <0.0001 | --- |
| GC × GC × GC | 0.013 | $(1.27,1.33) \times 10^{-2}$ | 7,263 | <0.0001 | --- |

(a) CI = Confidence Interval for the coefficient

transformation. However, this $R^2$ is substantially higher than that obtained with simple linear regression.

## 6.4 Evaluation of Regression Analysis Methods Based on Applications to the *E. coli* Model

In Sections 6-1 to 6-3 regression analysis was applied to different modules and parts of the *E. coli* model. Regression analysis was evaluated based upon applicability of the functional form of the model, the use of regression coefficients as an indicator of sensitivity, the use of confidence intervals for regression coefficients to evaluate the ambiguity of rankings, the use of F values as a sensitivity measure, and the ease of application.

The need to assume a specific functional relation between inputs and the output in a regression model is a disadvantage for this method. If the specific functional assumption is not comparable to the original model, the results from the regression analysis may not be valid. In these cases the fitted regression model addresses only a portion of the original model response variation. Estimated $R^2$ values of 0.82 to 0.90 in several case studies for the production module indicate that the linear assumption for the model response is a good approximation. In the slaughter module, the low $R^2$ values of 0.10 to 0.12 imply that there is not a linear relationship between inputs and the output. In the preparation module, the $R^2$ values for the growth estimation part were 0.51 to 0.52, indicating an approximately plausible goodness-of-fit for the linear assumption. In contrast, in the serving contamination part, the linear assumption for the relation between inputs and the output appeared to be poor, based upon $R^2$ values of 0.05 to 0.11.

Figure 6-21 depicts the scatter plot for the serving contamination in summer versus the grinder contamination. A linear regression model was fitted to data points in the scatter plot. The

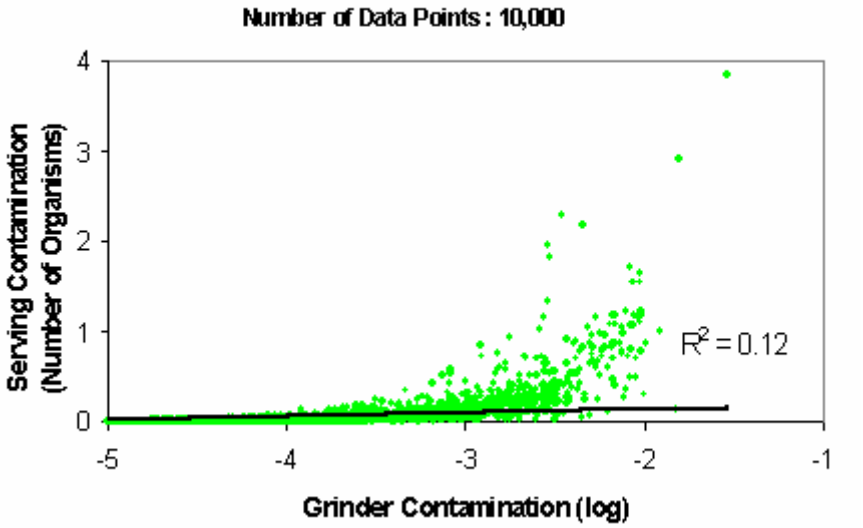



Figure 6-21. Comparison Between the Original Model Response in the Serving Contamination Part and the Linear Assumption.

$R^2$ coefficient is an index for evaluating the goodness-of-fit of the linear model to the data. An $R^2$ of 0.12 for this example indicates that the model explains only a small portion of the variability in the data. A key reason for the lack of good fit is the nonlinear response of the output to the input, which is not captured by the regression model. As complementary analyses in Section 6.3.5, rank regression, regression with log scale transformation, and regression with higher order terms were applied to selected case scenario in the serving contamination part in order to evaluate how these methods can improve the amount of variability in the output that can be captured by the fitted regression model. Moreover, comparison of the rankings based on the results obtained from these methods can give insight regarding the validity of the ranks when there is a low $R^2$ value for a simple linear regression analysis. Results of these analyses indicated that there was a substantial increase in the $R^2$ value using the rank regression or regression with log scale transformation. Fitted regression models in these cases captured almost all the output variation. There was also a noticeable increase in the $R^2$ value using higher order terms including interaction, quadratic, and cubic terms in the regression model. For the particular case study presented in Section 6.3.5, ranks of the inputs did not change when using different techniques for the regression analysis. Although the fitted linear regression model with low $R^2$ cannot be used for prediction purpose, this model can still be used for sensitivity analysis. This finding may

provide confidence in using the results of the sensitivity analysis based on linear regression even if the $R^2$ value is substantially low. But future works would be required to justify this finding in general. Practitioners should be cautioned that if the $R^2$ value is low, other methods should be used to confirm results or to develop more reliable results, such as alternative regression analysis using appropriate transformations (e.g., ranks, logarithmic transformations) or other techniques (e.g., ANOVA).

The growth estimation part was selected for application of the rank regression method to the two-dimensional simulation of variability under several uncertainty realizations. In this case, rank regression did not substantially improve the amount of variability captured by the fitted regression model. Although the equation used in the model for estimation of the growth is monotonic, there are some conditions in the model enforcing a constant growth even though there is variation in the model inputs. For example, no growth is estimated for cases where the storage time is less than the available lag period. These model characteristics prevent the model from responding in a completely monotonic pattern. Nonetheless, the rankings based on the rank regression method are somewhat different from that of the standardized linear regression method with respect to the most important inputs.

The use of regression coefficient estimates as a measure of sensitivity of the output to individual inputs was demonstrated in this chapter. Corresponding to each regression coefficient there is a standard error that can be used to derive the confidence interval for the coefficient. These confidence intervals for the coefficients can be used to evaluate the ambiguity of the ranks. The output has comparable sensitivity to inputs with overlapping confidence intervals for the estimated regression coefficients.

In a case with qualitative inputs in the model, coefficients are estimated for the indicator variables and not for the qualitative inputs. In order to compensate for this disadvantage, F values estimated for each input can be used as an index of sensitivity. Hence, the inputs can be ranked based upon the relative magnitude of the F values.

In Sections 6.3.2.1 and 6.3.2.2, Pearson and Spearman correlation coefficients methods were applied to the two-dimensional framework of variability and uncertainty in the growth estimation part. The relative magnitude of the correlation coefficient was presented as the measure of the sensitivity of the output to individual inputs. The assumption of linear association between the output and individual inputs is a potential disadvantage for the Pearson correlation

coefficients method. Although the Spearman correlation coefficients method for sensitivity analysis does not assume any linear functional relation in the model, this method assumes a monotonic relationship. Hence, for cases in which these assumptions are not satisfied, the results based upon these methods are not reliable. Moreover, neither the Pearson nor Spearman correlation coefficients method for sensitivity analysis can identify possible interaction effects in the model.

Regression analysis is a relatively easy to apply and interpret method for sensitivity analysis. In order to perform regression analysis, a dataset containing the values of each input in a Monte Carlo simulation and the corresponding values for each output of interest can be fed into any statistical software capable of performing regression analysis, such as SAS©. There are direct measures of sensitivity in the regression analysis that can be used for rank-ordering the inputs.

# 7 CLASSIFICATION AND REGRESSION TREES FOR THE *E. COLI O157:H7* MODEL

The purpose of this chapter is to apply CART to different modules and parts of the *E. coli* model for purpose of sensitivity analysis. CART is discussed in Section 2.2.3. The *E. coli* model is discussed in Chapter 3**.** Key advantages of CART are that it is possible to deal with both qualitative and quantitative inputs, to identify thresholds, and to gain insight into the sensitivity of inputs conditional on the values for other inputs. The importance of an input is indicated by whether it is selected as the basis for splitting the tree at the highest branches, and whether it is selected at multiple levels of the tree to further subdivide the data. The final nodes or leaves of the tree represent databases that have been created by a systematic partitioning of the data. The partitioned data under different nodes of the same branch have mean values that are significantly different from each other.

In order to gain additional insight into the sensitivity of model inputs conditional on partitioned data for different leaves of the regression tree, additional sensitivity analysis methods can be applied to such databases. For example, regression analysis or ANOVA can be applied for the purposes of determining which model inputs are most sensitive conditional on a particular partition of the original input data. One of the partitions will contain data that produces the largest mean value of the model output compared to other partitions of the data. To the extent that the results of complementary analyses on different nodal databases produce similar results, the ranking of sensitive inputs would be shown to be robust with respect to partitioning of the input data. However, it is more likely the case that the sensitivity analysis results will be different for different nodal databases. Such an outcome could, for example, help identify the combinations of input values that would produce the highest exposure or risk estimate. Therefore, in this work, CART is used as a first step followed by the application of regression analysis or ANOVA to nodal data.

A disadvantage of CART is that there is not a clear summary statistic via which to clearly rank the importance of different inputs. However, in Sections 7.3.1.1 and 7.3.1.2 a possible sensitivity index is explored and evaluated. Two case studies are provided in the growth estimation part to evaluate the amount of contribution of each selected input in the regression tree to the reduction of the total deviance as an alternative sensitivity index. Hence, inputs are ranked based on the percentage of their contribution to the reduction of the total deviance. This

alternative approach for ranking inputs is compared with rankings based on the visualization of the regression tree accompanied by complementary analyses.

In order to apply CART to different modules and parts of the *E. coli* food safety risk assessment model, S-PLUS© Version 6.1 was used. This software has the ability to perform CART analysis on a dataset, using a Graphical User Interface (GUI). Moreover, both qualitative and quantitative inputs can be addressed using the specific options of the software.

This chapter contains four parts, in which results are presented for the production module, slaughter module and preparation module in Sections 7.1, 7.2, and 7.3, respectively. Section 7.4 presents an evaluation of CART as a method for the sensitivity analysis. The limitations, advantages, disadvantages and key criteria for application of this method are summarized in this section.

## 7.1 Application of CART to the Production Module

In the production module CART is applied to four parts, including feedlot prevalence, within feedlot prevalence, breeding herd prevalence, and within breeding herd prevalence. The results of the analyses are presented in Sections 7.1.1 to 7.1.4 for these four parts, respectively.

### 7.1.1 Uncertainty in the Feedlot Prevalence Part

As explained in Section 3.2.1, for feedlot prevalence estimation, inputs include the apparent prevalence and the herd sensitivity as quantitative inputs, and the study as a qualitative one. Distributions for these inputs are summarized in Table 3-9. The output in this part is the median feedlot prevalence. For the feedlot prevalence part there is a one-dimensional uncertainty simulation with 65,000 iterations, as explained in Section 3.3.1.

In Figure 7-1 the result of CART analysis is depicted in the form of a regression tree. In this figure, *HS* stands for the herd sensitivity. Each level of the study is presented by a letter (e.g. *a, b*, *c*, and *d* for Dargatz & Hancock 1997, Hancock 1998, Smith 1999, and Elder 2000, respectively).

The regression tree in Figure 7-1 shows that the data for the feedlot prevalence is divided into two datasets based on studies. The first node in the regression tree indicates that if the study is *b* (i.e., Hancock 1998) the mean response is 0.49, as indicated at the bottom of the left-most branch of the tree. For other studies (i.e., Dargatz and Hancock 97, Smith 99, and Elder 2000) the dataset is further classified into two datasets based on the study *a* which is the Dargatz Hancock (1997) study and a separate partition based upon studies *c* and *d*, which are the Smith

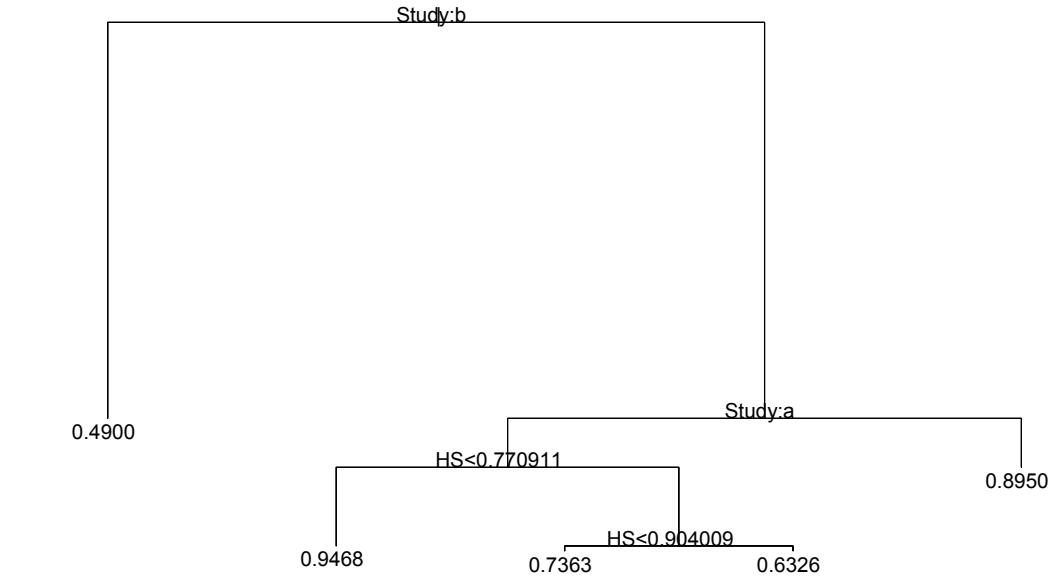



Figure 7-1. The Regression Tree for the Median Feedlot Prevalence.

(1999) and Elder (2000) studies, respectively. For the Dargatz and Hancock (1997) study the mean response depends on different values of the herd sensitivity, while for the other two studies the mean response is 0.895, as indicated at the bottom of the right-most branch of the tree. For the Dargatz & Hancock (1997) study the dataset is subdivided two times based upon the values of the herd sensitivity to create three partitions: (1) herd sensitivity less than 0.77 with a mean response of 0.947; (2) herd sensitivity between 0.77 and 0.90 with a mean response of 0.736; and (3) herd sensitivity greater than 0.90 with a mean response of 0.633.

In CART analysis no restriction was specified for the number of nodes in the regression tree. Hence, the mean responses presented in Figure 7-1 account for all of the variability in the output that could be captured by partitioning the dataset. Figure 7-1 indicates that the median feedlot prevalence is most sensitive to the study, because this input is placed at the first splitting node. Furthermore, the vertical distance below each split indicates the portion of deviance that is reduced because of the split. Thus, the long vertical distance below the first split for the study illustrates that separating study *b* from the other three studies accounts for most of the

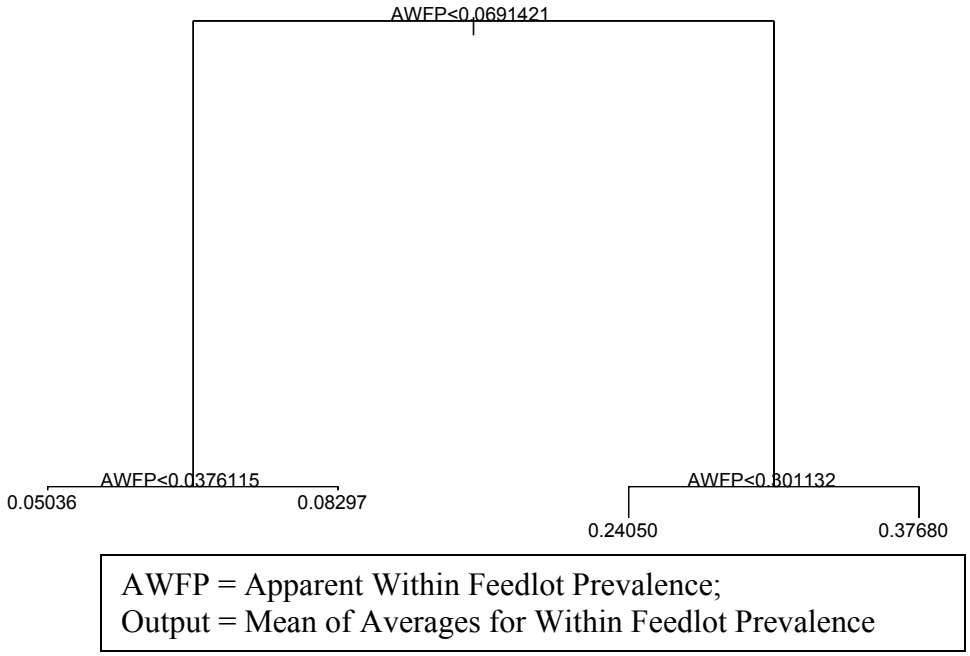



Figure 7-2. The Regression Tree for the Average Within Feedlot Prevalence.

explainable variability in the model output. Herd sensitivity has a rank of two, because it is placed in the lower nodes of the tree and under the splits based on the study. Moreover, mean values of the output range between 0.49 and 0.95 among the leaves, which illustrates that partitioning the data leads to approximately a factor of two difference in mean values at the leaves.

### 7.1.2 Uncertainty in the Within Feedlot Prevalence

Section 3.2.1 explains the within feedlot prevalence part of the production module. The inputs for this part include the apparent within feedlot prevalence and the test sensitivity as quantitative inputs, and the study and the season as qualitative inputs. Table 3-9 summarizes the distributions for these inputs. The output of interest is the average within feedlot prevalence. The case scenario for this part is based upon a one-dimensional uncertainty simulation with 65,000 iterations as described in Section 3.3.1.

In Figure 7-2 the result of CART analysis is depicted in the form of a regression tree. The first node in the regression tree subdivides the dataset into two divisions: (1) apparent within feedlot prevalence less than 0.069; and (2) apparent within feedlot prevalence greater than 0.069. The datasets corresponding to high and low values of the apparent within feedlot prevalence are

classified into four subdivisions based on values of 0.301 and 0.038 for the same input, respectively.

A high value of the response corresponds to high values of the apparent within feedlot prevalence and other inputs do not contribute to the explainable variability of the output. The highest mean response of 0.38 corresponds with the case that the apparent within feedlot prevalence is more than approximately 0.30. Furthermore, the long vertical distance below the first split for the apparent within feedlot prevalence illustrates that classifying the dataset based on the value of 0.069 of the apparent within feedlot prevalence accounts for most of the variability in the model output that can be captured by partitioning the data. Because only the apparent within feedlot prevalence was selected in the tree, the apparent within feedlot prevalence is ranked as the most important input.

### 7.1.3 Uncertainty in the Breeding Herd Prevalence Part

As described in Section 3.2.1, for breeding herd prevalence estimation, the inputs include the apparent prevalence and herd sensitivity as quantitative inputs, and study as a qualitative one. The output is median breeding herd prevalence. Distributions for the inputs are given in Table 3-9. The case scenario in Section 3.3.1 is based upon a one-dimensional uncertainty simulation with 65,000 iterations. In CART analysis for this part the number of leaves of the tree was specified as 7. Hence, in Figure 7-3 there are 7 mean responses presented in the regression tree. This number of nodes accounts for almost 95 percent of the variability in the output that can be captured if no restriction in the number of nodes was considered. The restriction on the size of the returned tree provides a more understandable tree with fewer splitting nodes and branches.

In Figure 7-3 the result of CART analysis is depicted in the form of a regression tree. Each level of the study is presented by a letter (i.e. *a, b*, *c*, *d, e, f* and *g* for Garber 1998, Sargeant 2000, Hancock/CFSAN 2001, Hancock 1997a, Hancock 1998, Lagreid 1998, and Hancock 1997b, respectively).

The regression tree shows that the data for the breeding herd prevalence is divided into two datasets based on the study. When the study is *a, b, c*, or *d*, the dataset is subdivided into four leaves based upon herd sensitivity and apparent prevalence. For other studies (i.e. *e, f,* or *g*) there is a split on the right side of the tree based on the study. If the study is *g,* the mean response depends on different values of herd sensitivity, while in cases that the study is *e* or *f* the mean response is 0.919, as indicated in the bottom of the right-most branch of the tree. The highest

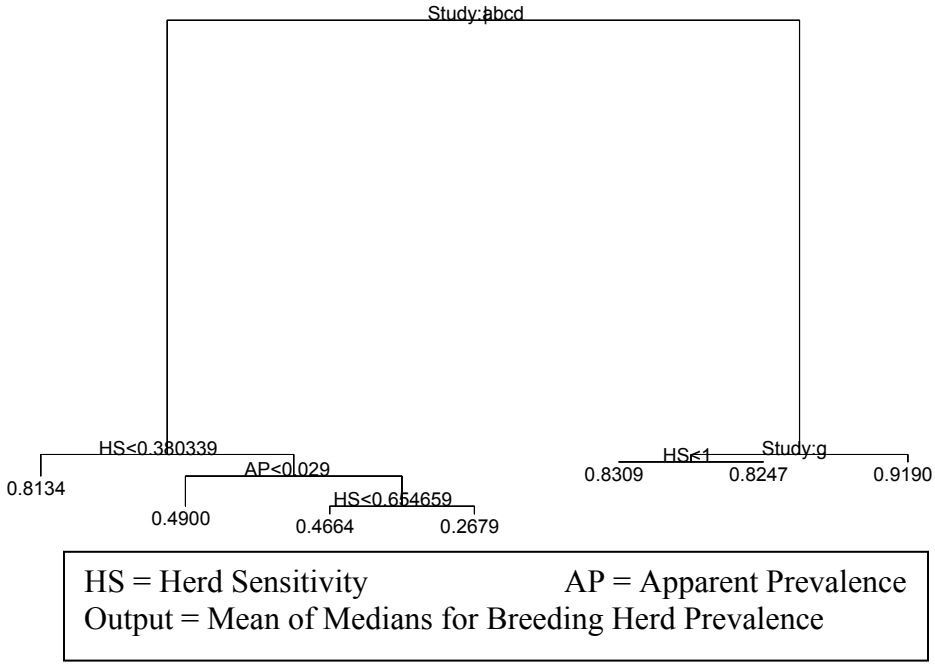



Figure 7-3. The Regression Tree for the Median Breeding Herd Prevalence.

mean response corresponds with the studies *e* or *f* with a value of 0.919. Furthermore, the long vertical distance below the first split for the study illustrates that classifying the dataset based on two sets of studies with *a, b, c,* and *d* in one group and *e, f,* and *g* in another group explains most of the variability that can be captured in the model output. Thus, the median breeding herd prevalence is most sensitive to the study.

The fact that study was selected as the first basis for splitting the data and that herd sensitivity was selected repeatedly throughout the lower notes of the tree suggest that these two inputs are important. Furthermore, it is clear from the tree that the highest values of median breeding herd prevalence, such as those mean values of 0.92 and higher beneath the nodes under the right branch of the tree, are associated with the use of specific studies. Therefore, to better understand or confirm which of the other inputs aside from study are important, the results of the CART analysis were supplemented with statistical sensitivity analysis methods applied to two datasets. One dataset was for model results associated with studies a, b, c, and d and the other was for model results based upon the other studies. The statistical method used was regression analysis. This method is explained in Sections 2.2.1, and was applied to the *E. coli* model as

described in Chapter 5. The results of the application of regression analysis to these two datasets are given in Table 7-1.

The rankings in Table 7-1 are based on the magnitude of the standardized regression coefficients. Rankings are presented for statistically significant inputs with Pr>F less than 0.05. The F values indicate that for the first dataset all inputs are statistically significant, while for the second dataset only study is statistically significant. Because study is a qualitative input, there is no coefficient estimate for this input. The results of the regression analysis imply that the herd sensitivity and apparent prevalence are ranked first and second, respectively. The 95 percentile confidence intervals are estimated in order to evaluate the robustness of the rankings. These intervals indicate that rankings are robust because there is no overlap.

Considering the results of the regression tree and the complementary sensitivity analysis, the study, herd sensitivity, and apparent prevalence are ranked first, second and third, respectively.

### 7.1.4 Uncertainty in the Within Breeding Herd Prevalence Part

Section 3.2.1 explains the within breeding herd prevalence part of the production module. Inputs in this part include the apparent within breeding herd prevalence and the test sensitivity as quantitative inputs, and the study and the season as qualitative inputs. Table 3-9 summarizes the distributions for these inputs. The output is the average within breeding herd prevalence. The case scenario for this part is based upon a one-dimensional uncertainty simulation with 65,000 iterations. In CART analysis for this part the number of leaves of the tree was specified as 7. Hence, in Figure 7-4 there are 7 mean responses presented in the regression tree. This number of nodes accounts for almost 85 percent of the variability in the output that can be captured if no restriction on the number of nodes was considered. The restriction on the size of the returned tree provides a more understandable tree with fewer splitting nodes and branches. In Figure 7-4, each level of the study is presented by a letter (i.e. *a, b*, *c*, *d, e,* and *f* for Garber 1998, Besser 1997, Rice 1997, Hancock 1994, Sargeant 2000, and Hancock/CFSAN 2001, respectively.

The regression tree in Figure 7-4 shows that the test sensitivity was selected for the first split and that this split accounts for a large reduction in deviance. The first node in the regression tree subdivides the dataset into two divisions: (1) test sensitivity less than 0.118; and (2) test sensitivity greater than or equal to 0.118. For cases with the test sensitivity less than 0.118, the dataset is subdivided further based upon test sensitivity as follows: (1) test sensitivity less than

Table 7-1. Regression Analysis Results for the Median Breeding Herd Prevalence ($R^2 = 0.65$ for the first dataset and $R^2 = 0.43$ for the second dataset)

| Variable | Coefficient | 95% CI [(a)] | F Value | Pr>F | Rank |
|---|---|---|---|---|---|
| **Garber 1998, Sargeant 2000, Hancock/CFSAN 2001, Hancock 1997a (a, b, c, and d)** | | | | | |
| Study | --- | --- | 6,500 | <0.0001 | --- |
| Apparent Prevalence | -0.13 | (-0.14, -0.12) | 350 | <0.0001 | 2 |
| Herd Sensitivity | -0.44 | (-0.45, -0.44) | 35,000 | <0.0001 | 1 |
| **Hancock 1998, Lagreid 1998, Hancock 1997b (e, f, and g)** | | | | | |
| Study | --- | --- | 10,000 | <0.0001 | --- |
| Apparent Prevalence | 0.26 | (0.16, 0.36) | 1 | 0.54 | --- |
| Herd Sensitivity | -0.001 | $(-5,3)\times10^{-3}$ | 1.5 | 0.25 | --- |

(a) CI = Confidence Interval for the coefficient

0.06 with a mean response of 0.88; and (2) test sensitivity between 0.06 and 0.118 with a mean response of 0.55. For high values of the test sensitivity, presented in the right side of the regression tree, the data are subdivided based on study and then further divided based upon apparent breeding herd prevalence and test sensitivity. The highest average within breeding herd prevalence of 0.88 corresponds to cases where test sensitivity is less than 0.06.

The fact that test sensitivity was selected as the first basis for splitting the data and that it was selected repeatedly throughout the lower nodes of the tree suggest that this input is important. Therefore, to better understand or confirm which of the other inputs are important, the results of the CART analysis were supplemented with statistical sensitivity analysis methods applied to two datasets classified based on test sensitivity. Thus, ANOVA was used as a complementary sensitivity analysis in order to rank the inputs conditional on test sensitivity. The first dataset includes data with test sensitivity less than 0.117 and the second dataset contains data with test sensitivity greater than or equal to 0.117. The results of the complementary analyses are given in Table 7-2.

The rankings in Table 7-2 are based on the magnitude of the F values. Rankings are presented for statistically significant inputs with Pr>F less than 0.05. For both datasets all inputs are statistically significant. The magnitudes of the F values indicate that in both datasets study, apparent within breeding herd prevalence, and season are ranked first, second, and third, respectively. The large difference between the F values for statistically significant inputs indicates that the rankings are robust. For example, the F value for study is approximately 3.9

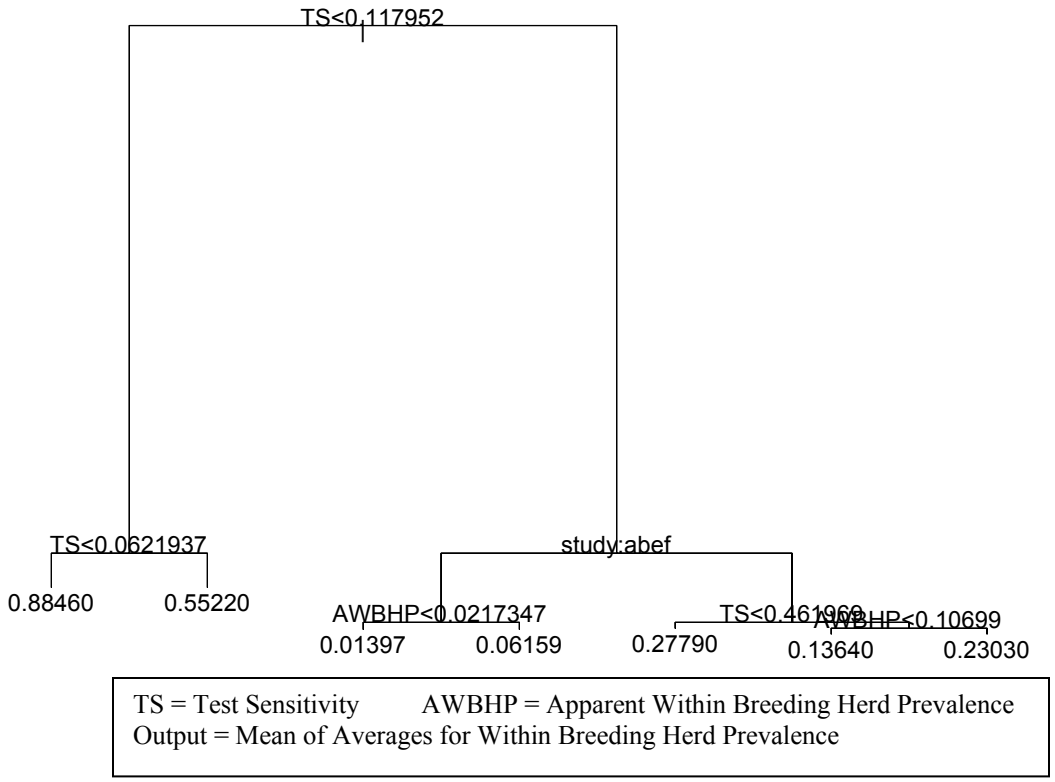



Figure 7-4. The Regression Tree for the Average Within Breeding Herd Prevalence.

Table 7-2. ANOVA Results for the Median Breeding Herd Prevalence

| Variable | F Value | Pr > F | Significant | Rank |
|---|---|---|---|---|
| **Test Sensitivity < 0.117** | | | | |
| Study | 4,250 | <0.0001 | Yes | 1 |
| Season | 12 | 0.0005 | Yes | 3 |
| Apparent Within Breeding Herd Prevalence | 1,100 | <0.0001 | Yes | 2 |
| **Test Sensitivity >= 0.117** | | | | |
| Study | 8,830 | <0.0001 | Yes | 1 |
| Season | 23 | <0.0001 | Yes | 3 |
| Apparent Within Breeding Herd Prevalence | 3,230 | <0.0001 | Yes | 2 |

and 2.8 times greater than the F value for the apparent within breeding herd prevalence in the first and second dataset, respectively.

Considering the regression tree and results of the complementary sensitivity analysis, test sensitivity, study, apparent within breeding herd prevalence, and season are ranked first, second, third, and fourth, respectively.

## 7.2 Application of CART to the Slaughter Module

The slaughter module is discussed in Section 3.2.2. Inputs and corresponding distributions in the slaughter module are summarized in Table 3-10. The output of interest in the slaughter module is the contamination in combo bins. Three different types of probabilistic analysis were performed for this module, as described in Section 3.3.2: (1) one-dimensional simulation of variability based upon mean values of uncertain inputs; (2) two-dimensional simulation of variability for each realization of uncertainty; and (3) one-dimensional simulation of both variability and uncertainty co-mingled.

In CART analysis, since inputs are ranked based on visual inferences from the regression tree, and in some cases by incorporation of complementary sensitivity analysis, application of CART for the second simulation considering variability for several uncertainty realizations was impractical. Thus, in the slaughter module, CART analysis was only applied to the first and third analyses. The results of CART analysis for variability only and one-dimensional co-mingled variability and uncertainty simulations are presented in Sections 7.2.1 and 7.2.2, respectively. Moreover, Section 7.2.3 compares the results from Sections 7.2.1 and 7.2.2.

### 7.2.1 Variability Only

This section presents the results of CART analysis applied to a one dimensional probabilistic simulation in which variability is only considered for mean uncertainty, based upon the case study scenario described in Section 3.3.2. The results of CART analysis are depicted in Figure 7-5 in the form of a regression tree.

The regression tree in Figure 7-5 shows that the data were divided into two datasets based upon the chilling effect. The first node in the regression tree subdivides the dataset into two divisions as follows: (1) chilling effect less than 0.39 logs; and (2) chilling effect greater than or equal to 0.39 logs. For the former case, the mean response is 0.51 as presented in the left-most terminal node of the tree. For the latter case, the dataset is subdivided twice considering values of the number of organisms and chilling effect. Thus, there are three mean responses based on the

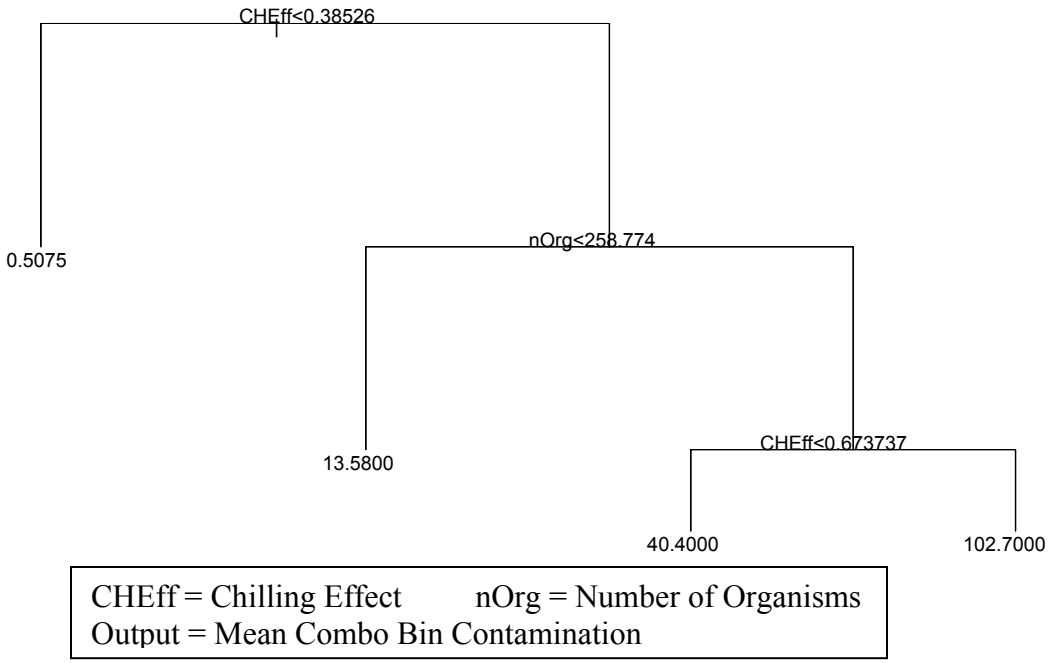



Figure 7-5. The Regression Tree for the Combo Bin Contamination from Steer and Heifer in Summer for the Variability only Analysis.

values of these inputs as follows: (1) chilling effect greater than or equal to 0.39 logs and the initial number of organism less than 259 with a mean response of 13.58; (2) chilling effect between 0.39 and 0.67 logs and number of organism greater than or equal to 259 with a mean response of 40.4; and (3) chilling effect greater than or equal to 0.67 logs and number of organism greater than or equal to 259 with a mean response of 103. The regression tree implies that high values of combo bin contamination (i.e. 103 *E. coli* organisms or 2 logs of contamination) correspond with cases in which the chilling effect is more than an approximate value of 0.67 logs and number of organisms is greater than 259.

In CART analysis no restriction was specified for the number of nodes in the regression tree. Hence, the mean responses presented in Figure 7-5 account for all of the variability in the output that could be captured by partitioning the dataset. The regression tree implies that the combo bin contamination is most sensitive to the chilling effect, because this input is placed at the first node of the tree. Hence, the chilling effect is ranked first. The number of organisms is ranked second. Other inputs were not selected in the regression tree by CART analysis.

Therefore, the other inputs are likely to be less sensitive than the ones that were selected in the tree.

### 7.2.2 One-Dimensional Simulation of Variability and Uncertainty

This section presents the results of CART analysis applied to a one dimensional probabilistic simulation in which variability and uncertainty are co-mingled, based upon the case study scenario described in Section 3.3.2. The results of CART analysis are depicted in Figure 7-6 in the form of a regression tree. In CART analysis for this module, the maximum number of leaves of the tree is specified as 8. Hence, in Figure 7-6 there are 8 mean responses presented in the regression tree. This number of nodes account for almost 90 percent of the variability in the output that can be captured if no restriction on the number of nodes was considered.

The regression tree in Figure 7-6 shows that the data was divided into two datasets based upon the chilling effect. The first node in the regression tree subdivides the dataset into two divisions as follows: (1) chilling effect less than 2.2 logs; and (2) chilling effect greater than or equal to 2.2 logs. These datasets are further subdivided in the left and right branches of the regression tree using values for inputs such as number of organisms, chilling effect, and washing effect for the left branch, and number of organisms and contaminated $cm^2$ of meat trim in the right branch. The highest combo bin contamination corresponds with cases in which the initial number of organisms on carcasses is greater than approximately of 128 organisms and the chilling effect is higher than 2.2 logs. The mean response for these cases is 3641 *E. coli* organisms or approximately 3.6 logs of contamination. The mean responses range between 3.9 and 3640 *E. coli* organisms per combo bin.

The long vertical distance below the second right split for the number of organisms illustrates that classifying the dataset based on a value of 127.9 for the number of organisms when the chilling effect is greater than or equal to 2.2 logs explains most of the variability that can be captured in the model output. The regression tree implies that the combo bin contamination is most sensitive to both the chilling effect and the number of organisms. The chilling effect is placed at the first node of the tree. However, although the number of organisms was not selected until the second node in the right-most branch, this input discriminates the mean response of 3641 from other leaves with mean responses of 130 to 528. Thus, the partitioning of data for large values of chilling effect with respect to the number of organisms accounts for a wide range of variation in the response. Therefore, there appears to be an important interaction

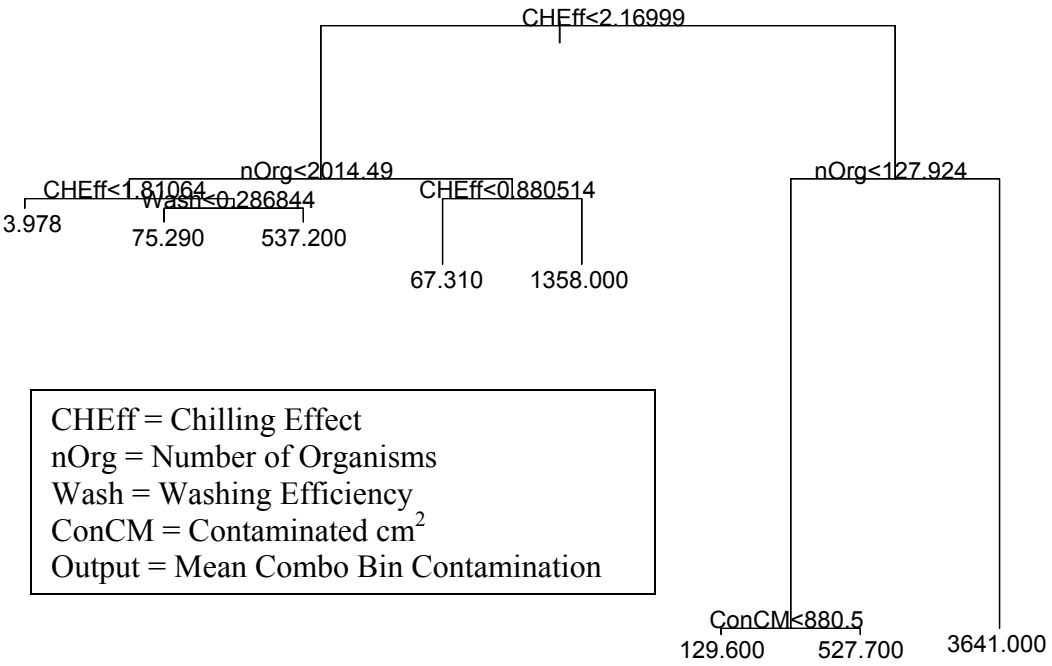



Figure 7-6. The Regression Tree for the Combo Bin Contamination from Steer and Heifer in Summer for One-Dimensional Variability and Uncertainty Analysis.

between chilling effect and number of organisms. For low values of chilling effect, the mean response varies from 4.0 to 1358 depending on more refined ranges of chilling effect, number of organisms, and washing efficiency. For high values of chilling effect, the number of organisms is the most important input. Therefore, it may be the case that chilling effect and number of organisms are of comparable importance. The other inputs selected in the tree are of minor importance. Inputs not selected in the tree are deemed to be unimportant. To gain further insight into which inputs are important conditional on the chilling effect, regression analysis was applied for two cases.

Table 7-3 summarizes the results of the complementary regression analysis applied to the dataset with chilling effect of less than 2.2 logs. The inputs in Table 7-4 are ranked based on the magnitude of regression coefficients. These rankings are only presented for statistically significant inputs with Pr>F less than 0.05. The F values indicate that there is no statistically

Table 7-3. Regression Analysis Results for the Dataset with Chilling Effect Less than 2.2 logs ($R^2 = 0.11$)

| Variable | Coefficient | 95% CI [(a)] | *F Value* | Pr>F | Rank |
|---|---|---|---|---|---|
| Total Number of Combo Bin for Each Carcass | -0.010 | (-0.018,-0.003) | 7 | 0.008 | 11 |
| Total Number of Infected Animals | -0.030 | (-0.045, -0.019) | 23 | <0.0001 | 7 |
| Total Number of Contaminated Animals | -0.024 | (-0.038, -0.010) | 11 | 0.009 | 9 |
| Probability of Positive Cases at both Steps of Dehiding and Evisceration | -0.032 | (-0.041, -0.024) | 59 | <0.0001 | 6 |
| Number of Positive Cases at both Steps of Dehiding and Evisceration | 0.012 | (0.002,0.022) | 6 | 0.018 | 10 |
| Number of Positive Cases at Evisceration | -0.008 | (-0.018, -0.002) | 2 | 0.13 | --- |
| Chilling Effect | 0.157 | (0.145,0.169) | 622 | <0.0001 | 2 |
| Number of Organisms | 0.235 | (0.226,0.243) | 2865 | <0.0001 | 1 |
| Trim/Vacuum/Washing Efficiency | -0.137 | (-0.149, -0.125) | 506 | <0.0001 | 3 |
| Evisceration Organisms Added | 0.029 | (0.021,0.036) | 59 | <0.0001 | 8 |
| Washing Effect | 0.110 | (0.10,0.12) | 572 | <0.0001 | 4 |
| Contaminated $cm^2$ | 0.100 | (0.09,0.11) | 359 | <0.0001 | 5 |

significant influence for the number of positive cases at evisceration. Therefore, the output is not sensitive to the variability in this input.

Based upon the magnitude of the coefficients for the statistically significant inputs, number of organisms, chilling effect, Trim/Vacuum/Washing efficiency, and washing effect arethe most sensitive inputs. In order to evaluate the robustness of the estimated rankings, the 95 percentile confidence intervals are estimated for statistically significant coefficients. Estimated confidence intervals for regression coefficients indicate that the rank of the number of organisms is robust. The confidence interval for this input does not overlap with the confidence interval of the second ranked input. In contrast, the ranks for the second and third important inputs are not robust, because their confidence intervals overlap.

Table 7-4. Regression Analysis Results for the Dataset with Chilling Effect Greater than or Equal to 2.2 logs ($R^2 = 0.54$)

| Variable | Coefficient | 95% CI [(a)] | *F Value* | Pr>F | Rank |
|---|---|---|---|---|---|
| Total Number of Combo Bin for Each Carcass | 0.089 | (0.035,0.140) | 11 | <0.0001 | 9 |
| Total Number of Infected Animals | 0.125 | (0.05,0.20) | 11 | <0.0001 | 7 |
| Total Number of Contaminated Animals | 0.180 | (0.092,0.269) | 16 | <0.0001 | 6 |
| Probability of Positive Cases at both Steps of Dehiding and Evisceration | 0.016 | (-0.036,0.069) | 0.4 | 0.7 | --- |
| Number of Positive Cases at both Steps of Dehiding and Evisceration | -0.117 | (-0.196, -0.037) | 9 | 0.0008 | 8 |
| Number of Positive Cases at Evisceration | 0.188 | (0.107,0.269) | 21 | <0.0001 | 5 |
| Chilling Effect | 0.406 | (0.33,0.48) | 114 | <0.0001 | 2 |
| Number of Organisms | 0.338 | (0.276,0.400) | 114 | <0.0001 | 3 |
| Trim/Vacuum/Washing Efficiency | 0.229 | (0.161,0.296) | 45 | <0.0001 | 5 |
| Evisceration Organisms Added | 0.234 | (0.179,0.289) | 71 | <0.0001 | 4 |
| Washing Effect | 0.450 | (0.338,0.512) | 200 | <0.0001 | 1 |
| Contaminated $cm^2$ | -0.012 | (0.086,0.062) | 0.1 | 0.8 | 10 |

Table 7-4 summarizes the results of the complementary regression analysis applied to the dataset with chilling effect greater than or equal to 2.2 logs. The inputs in Table 7-5 are ranked based on the magnitude of regression coefficients. These rankings are only presented for statistically significant inputs with Pr>F less than 0.05. The F values indicate that there is no statistically significant influence of the probability of positive cases at both steps of dehiding and evisceration. Therefore, the output is not sensitive to the variability in this input.

Based upon the magnitude of the coefficients for the statistically significant inputs, washing effect, chilling effect, number of organisms, and evisceration organisms added are the most sensitive inputs. In order to evaluate the robustness of the estimated rankings, the 95 percent confidence intervals are estimated for statistically significant coefficients. There is overlap of confidence intervals for regression coefficients in some cases. For example, the top

Table 7-5. Summary of the CART Analysis Results Based on Variability Only and Co-mingled Variability and Uncertainty Analyses

| Variable | Ranks | |
|---|---|---|
| | **Analysis 1[1]** | **Analysis 3[2]** |
| Total Number of Combo Bins for Each Carcass | NS[3] | 11[4] |
| Total Number of Infected Animals | NS[3] | 7[4] |
| Total Number of Contaminated Animals | NS[3] | 9[4] |
| Probability of Positive Cases at both Steps of Dehiding and Evisceration | NS[3] | 8[4] |
| Number of Positive Cases at both Steps of Dehiding and Evisceration | NS[3] | 10[4] |
| Number of Positive Cases at Evisceration | NS[3] | NA[5] |
| Chilling Effect | 1 | 1 |
| Number of Organisms | 2 | 2 |
| Trim/Vacuum/Washing Efficiency | NS[3] | 3[4] |
| Evisceration Organisms Added | NS[3] | 5[4] |
| Washing Effect | NS[3] | 4[4] |
| Contaminated $cm^2$ | NS[3] | 6[4] |

(1) Ranks based on the variability only analysis.
(2) Ranks based on the one-dimensional co-mingled variability and uncertainty analysis.
(3) NS = Not selected in the regression tree.
(4) Ranked based upon the complementary regression analysis.
(5) NA = Rank not available in the complementary regression analysis.

ranked input is significantly more important than the fourth and lower ranked inputs, but the second and third ranked input could be of comparable importance to the first ranked input. Thus, there is some ambiguity in the rankings.

The results for the high and low partitions of the data based upon the chilling effect were qualitatively similar. In both cases, chilling effect, number of organisms, Trim/Vacuum/Washing efficiency, and the washing efficiency were identified as a group of top four important inputs. Since chilling effect was selected by CART as the first basis for subdividing the data, chilling effect, number of organisms, washing effect, and Trim/Vacuum/Washing efficiency are deemed to be the most important inputs.

### 7.2.3 Summary and Comparison of the Results of CART Analysis in the Slaughter Module

In Sections 7.2.1 and 7.2.2, CART analysis was applied to two datasets considering variability only and co-mingled variability and uncertainty in inputs, respectively. In this section rankings based on these analyses are summarized and compared. Table 7-5 gives the ranks for each input based on analyses in Sections 7.2.1 and 7.2.2.

CART analysis identified the chilling effect and number of organisms on contaminated carcasses as the most important inputs for both the variability only and one-dimensional co-mingled variability and uncertainty analyses. The regression tree in the former analysis did not select other inputs. For the latter case, regression analysis was used as a complementary method for ranking the input considering information from the regression tree. Thus, more inputs were ranked in this case. The results were similar for the top two inputs based upon both probabilistic simulation approaches.

## 7.3 Application of CART to the Preparation Module

In the preparation module CART analysis was applied to three parts: (1)growth estimation; (2) cooking effect; and (3)serving contamination part. The results of the analyses are presented in Sections 7.3.1 to 7.3.3 for each of these three parts, respectively.

### 7.3.1 Application of CART to the Growth Estimation Part

The growth estimation part is discussed in Section 3.2.3. Three different types of probabilistic analysis were performed for this part, as described in Section 3.3.3: (1) one-dimensional simulation of variability based upon mean values of uncertain inputs; (2) two-dimensional simulation of variability for each realization of uncertainty; and (3) one-dimensional simulation of both variability and uncertainty co-mingled.

In the growth estimation part CART analysis was only applied to the first and third analyses. The results of CART analysis for variability only and one-dimensional co-mingled variability and uncertainty simulations are presented in Sections 7.3.1.1 and 7.3.1.2, respectively. Moreover, Section 7.3.1.3 compares the results from Sections 7.3.1.1 and 7.3.1.2.

Two case studies are provided in the growth estimation part in which a new sensitivity index is presented for CART. As presented in previous sections, the ranking of the inputs in CART is based on visualization of the regression tree and judgment. Complementary analyses on sub-divided datasets with other sensitivity analysis methods such as regression and ANOVA can be also informative. As explained in Section 2.2.3, CART reduces the total deviance in the dataset by subdividing the original dataset into more homogeneous subgroups. Hence, for each input selected at splitting nodes there is an associated reduction in the total deviance. Therefore, inputs selected in the regression tree can be ranked based on their contribution to the amount of reduction of the total deviance. In order to use this sensitivity index no limitation should be

considered in the total number of nodes in the regression tree, in order to achieve the maximum possible total reduction in deviance.

#### 7.3.1.1 Variability Only

This section presents the results of CART analysis applied to a one-dimensional probabilistic simulation in which variability is only considered for mean uncertainty, based upon the case study scenario described in Section 3.3.3. The results of CART analysis are depicted in Figure 7-7 in the form of a regression tree.

In CART analysis for this part the maximum number of leaves of the tree was specified as 10. Hence, in Figure 7-7 there are 10 mean responses presented in the regression tree. This number of nodes account for almost 85 percent of the variability in the output that can be captured if no restriction in the number of nodes was considered. The restriction in the size of the returned tree provides a more understandable tree with fewer splitting nodes and branches. The temperature at stage 3 was the first input selected as the basis for partitioning data. The time at stage 3 was selected at the second node under both branches of the first node. The temperature at stage 3 and the time at stage 3 appear in some of the subordinate nodes. These two inputs alone result in partitions of the data set with mean responses that vary from 0.07 to 0.52, as represented by the four leaves on the right of the tree. For stage 3 storage temperatures of less than 14.7 $^{o}$C and storage times of less than 59 hours, several other inputs were selected to partition the data as shown in the six left-most leaves of the tree. The mean response in these six leaves varies from 0.008 to 0.47. The mean response of 0.47 is associated with a stage 3 storage temperature of less than 14.7 $^{o}$C, a stage 3 storage time of less than 59 hours, a stage 1 storage time of greater than 72 hours, and a stage 1 storage temperature of greater than 10 $^{o}$C. This implies that a long storage time and a high storage temperature in stage 1 can lead to large growth even if the storage time and temperature in stage 3 are kept low. Conversely, as indicated by the right most leaf of the tree that has an average response of 0.52, if the storage time and temperature of stage 3 are greater than 14.7 $^{o}$C and 19 hours, respectively, then on average there is high growth irrespective of the time and temperature history of other stages. Another high growth scenario with an average response of 0.45 is based upon a stage 3 storage temperatures between 10.5 and 14.7 $^{o}$C for stage 3 storage times of greater than 59 hours.

The results for the three nodes with average responses of 0.47 to 0.52 illustrate that there are different ways in which high growth can occur. Therefore, these results illustrate that the

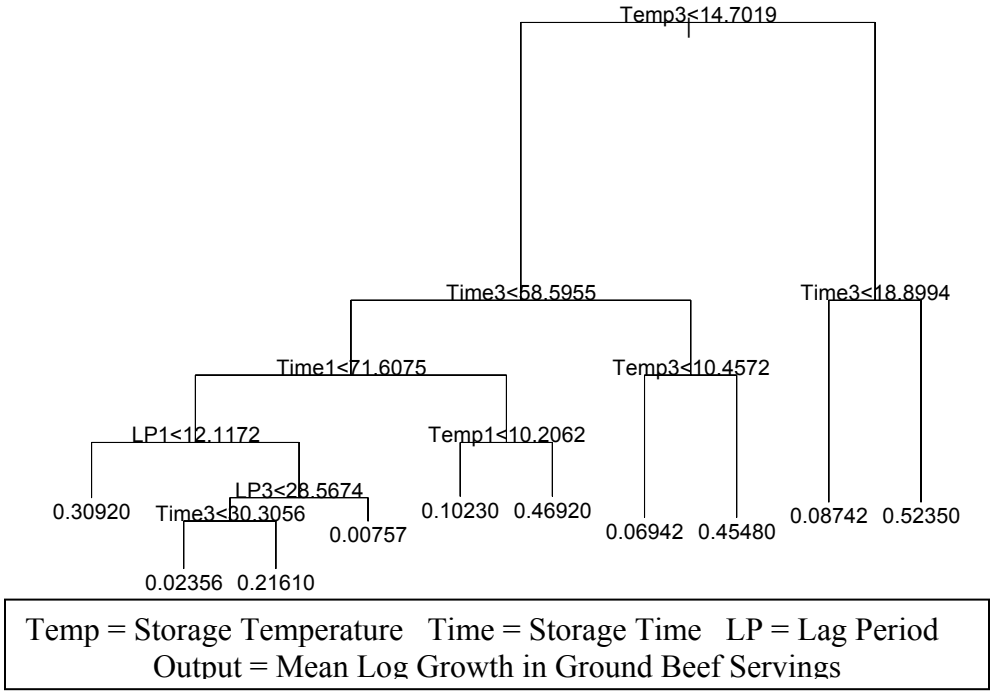



Figure 7-7. The Regression Tree for the Growth Estimation Part for the Variability Only Analysis.

regression tree is able to reflect complex interactions among the input values. These interactions occur because of nonlinear responses of growth rates to different combinations of time and temperature for different stages. Furthermore, because growth rates are small for small temperatures and/or short storage times, there are practical thresholds below which growth is comparatively insignificant. Thus, the results of CART analysis provide insight into the possible existence of such nonlinearities and thresholds.

The largest mean values of the response are associated with specific combinations of stage 3 storage time and temperature and are also influenced by the stage 1 storage time and temperature. Thus, these four inputs collectively appear to comprise the most important group of inputs. Other inputs, such as lag period at stages 1 and 3 ($LP_1$ and $LP_3$), were selected in the bottom left-most nodes of the tree. However, these inputs discriminate among mean responses of 0.007 to 0.31, which are considerable smaller than those in the range of 0.47 to 0.52 described

above. Thus, these latter two inputs are not as important as the first four. The response is likely to be insensitive to inputs that were not selected in the tree.

In order to further explore the sensitivity of the response to different inputs, regression analysis was performed for data partitioned based upon a stage 3 storage temperature of 14.7 $^{o}$C. The regression analysis results are summarized in Table 7-6.

In Table 7-6, the inputs are ranked based on the magnitude of standardized regression coefficients. Rankings are presented for statistically significant inputs with Pr>F less than 0.05. Based upon the F values, storage temperature, lag period, and generation time at stage 2 when the storage time at stage 3 is less than 14.7 $^{o}$C are not statistically significant. When the storage time at stage 3 is greater than or equal to 14.7 $^{o}$C, storage temperature and storage time at stage 2, lag period and generation time at stages 1 and 2, and storage temperature at stage3 are not statistically significant.

Based upon the magnitude of the coefficients for the statistically significant inputs when the storage temperature at stage 3 is less than 14.7 $^{o}$C, the storage temperature at stage 1, the storage time at stage 3, the storage temperature at stage 3, and the storage time at stage 1 are the top four important inputs. In order to evaluate the robustness of the estimated rankings, the 95 percent confidence intervals are estimated for coefficients. Estimated confidence intervals for regression coefficients indicate that the rankings for the storage temperature at stage 1 is robust, because the confidence interval for this input does not overlap with the confidence intervals of the worst ranked inputs. The ranks for the other top three inputs are not robust. The confidence intervals for the second and third inputs and for the third and fourth inputs overlap indicating that they may be of comparable importance.

When the storage temperature at stage 3 is greater than or equal to 14.7 $^{o}$C, the storage time at stage 3, the storage temperature at stage 1, the lag period and the generation time at stage 3 are top four important inputs. The 95 percent confidence intervals are estimated for statistically significant coefficients to evaluate the robustness of rankings. Estimated confidence intervals for regression coefficients indicate that the rankings for the top input is robust, because the confidence intervals for this input and the second ranked input do not overlap. The inputs ranked second to fourth do not have robust rankings, because their confidence intervals overlap. Hence, these inputs are of comparable importance.

Table 7-6. Regression Analysis Results for the Growth Estimation Part for Variability Only Analysis ($R^2$ = 0.45 for the first dataset and $R^2$ = 0.72 for the second dataset)

| Variable | Coefficient | 95% CI [(a)] | F Value | Pr>F | Rank |
|---|---|---|---|---|---|
| **Storage Temperature, Stage 3 <14.7$^{oC}$** | | | | | |
| Storage Temperature, Stage 1 | 0.42 | (0.41,0.43) | 8,200 | <0.0001 | 1 |
| Storage Temperature, Stage 2 | 0.01 | (-0.01,0.03) | 1 | 0.2 | --- |
| Storage Temperature, Stage 3 | 0.34 | (0.33,0.35) | 3,400 | <0.0001 | 3 |
| Storage Time, Stage 1 | 0.325 | (0.32,0.33) | 12,400 | <0.0001 | 4 |
| Storage Time, Stage 2 | 0.01 | (0.0,0.02) | 5 | 0.002 | 9 |
| Storage Time, Stage 3 | 0.35 | (0.347,0.353) | 14,150 | <0.0001 | 2 |
| Maximum Density | 0.01 | (0.004,0.016) | 10 | <0.0001 | 9 |
| Lag Period, Stage 1 | -0.02 | (-0.03,-0.01) | 20 | <0.0001 | 8 |
| Lag Period, Stage 2 | -0.004 | (-0.013,0.005) | 1 | 0.2 | --- |
| Lag Period, Stage 3 | -0.03 | (-0.04,-0.02) | 47 | <0.0001 | 7 |
| Generation Time, Stage 1 | 0.11 | (0.10,0.12) | 540 | <0.0001 | 5 |
| Generation Time, Stage 2 | 0.002 | (-0.009,0.012) | 0.1 | 0.8 | --- |
| Generation Time, Stage 3 | 0.06 | (0.05,0.07) | 134 | <0.0001 | 6 |
| **Storage Temperature, Stage 3 >=14.7$^{oC}$** | | | | | |
| Storage Temperature, Stage 1 | 0.19 | (0.16,0.22) | 132 | <0.0001 | 2 |
| Storage Temperature, Stage 2 | 0.01 | (-0.03,0.05) | 0.2 | 0.7 | --- |
| Storage Temperature, Stage 3 | 0.014 | (-0.032,0.06) | 0.4 | 0.6 | --- |
| Storage Time, Stage 1 | 0.12 | (0.10,0.14) | 137 | <0.0001 | 5 |
| Storage Time, Stage 2 | -0.016 | (-0.036,0.005) | 2 | 0.1 | --- |
| Storage Time, Stage 3 | 0.77 | (0.75,0.79) | 5,300 | <0.0001 | 1 |
| Maximum Density | 0.09 | (0.07,0.11) | 73 | <0.0001 | 6 |
| Lag Period, Stage 1 | 0 | (-0.03,0.03) | 0 | 0.9 | --- |
| Lag Period, Stage 2 | -0.003 | (-0.04,0.03) | 0 | 0.9 | --- |
| Lag Period, Stage 3 | -0.136 | (-0.17,-0.11) | 71 | <0.0001 | 3 |
| Generation Time, Stage 1 | 0.03 | (0.0,0.06) | 3 | 0.06 | --- |
| Generation Time, Stage 2 | 0.008 | (-0.031,0.046) | 0.2 | 0.7 | --- |
| Generation Time, Stage 3 | -0.13 | (-0.17,-0.09) | 40 | <0.0001 | 4 |

(a) CI = Confidence Interval for the coefficient.

Based upon the results from both CART and the complementary regression analysis, a judgment was made that the most important inputs are ranked in the following order, starting with the most important: (1) stage 3 storage temperature; (2) stage 3 storage time; (3) stage 1storage temperature; and (4) stage 1 storage time. The regression results confirmed that the stage 1 storage temperature and time were important inputs, conditional on specific ranges of the stage 3 storage temperature.

The amount of contribution of each input to the reduction of the total deviance is considered as an alternative sensitivity index. The dataset for the variability only analysis in the growth estimation part has a total deviance of 807. If no condition is considered for the number of nodes, the regression tree can capture 84 percent of the deviance (i.e., 677.3). Table 7-7 summarizes the contribution of each input to the reduction of total deviance. Eight inputs were selected in the regression tree. These inputs include storage times at stages 1 and 3, storage temperatures at stages 1 and 3, maximum density, lag period at stage 3, and generation time at stage 3. Table 7-7 indicates that there were 7 levels in the regression tree. Except for the first level of the tree, 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 this table. Storage temperature at stage 3 was selected in the first split of the tree. Dividing the dataset based on the condition provided for the storage temperature at stage 3 in the first splitting node reduced the total deviance by approximately 20 percent. Storage time at stage 3 was selected twice in the second level of the tree. Selection of storage time at stage 3 in the second level led to approximately 20 percent reduction in the total deviance. At the seventh level of the tree, three inputs were selected. Storage time at stage 1 was selected twice, while storage time at stage 3 and lag period at stage 1 were selected once. Selection of these inputs at this level leaded to approximately 0.7, 3.5, and 1.0 percent reduction in the total deviance, respectively.

For each input in Table 7-7 the percent of contribution to the total reduction in the deviance is identified. These contributions vary between 0.5 and 36.5 percent. Selected inputs in the regression tree are ranked based on their contribution to the total deviance reduction. The ranking indicates that storage temperature at stage 3, storage time at stage 3, and storage time at stage 1 were selected as the top three inputs to the model. Based on the rankings the input

Table 7-7. Reduction in Deviance Associated with Selected Inputs in the Regression Tree Generated in the Growth Estimation Part for the Variability Only Analysis

| Level of the Tree | Selected Inputs in the Regression Tree (1) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | **$Time_1$** | **$Time_3$** | **$Temp_1$** | **$Temp_3$** | **MD** | **$LP_1$** | **$LP_3$** | **$GT_3$** |
| **1st Level** | | | | 163.3 | | | | |
| **2nd Level** | | 44.1<br>118.7 | | | | | | |
| **3rd Level** | 39.7 | 7.3<br>13.7 | | 83.6 | | | | |
| **4th Level** | | 9.8 | 31.6 | | 3.2 | 32.3 | | 7.3<br>8.6 |
| **5th Level** | 14 | | | | | 7.0 | 13.8<br>6.3 | 2.9 |
| **6th Level** | 3.5<br>2.1 | 28.0 | | | | 7.9 | | |
| **7th Level** | 16.7<br>2.9 | 3.6 | | | | 5.6 | | |
| **Sum** | 78.8 | 44.1 | 31.6 | 246.9 | 3.2 | 52.7 | 20.1 | 18.8 |
| **Percent of Contribution (2)** | 11.6 | 33.2 | 4.7 | 36.5 | 0.5 | 7.8 | 3.0 | 2.8 |
| **Rank** | 3 | 2 | 5 | 1 | 8 | 4 | 6 | 7 |

(1) Time = Storage Time, Temp = Storage Temperature, MD = Maximum Density, LP = Lag period, and GT = Generation Time. Subscript numbers indicate corresponding stage of the growth process.
(2) Total deviance of the dataset is 807. The amount of deviance captured by the regression tree is 677.2.

selected at the first splitting node has the highest contribution to reduction of the total deviance. Moreover, except the input selected in the first splitting node, the greater the number of times each input is selected in the tree, the higher is the corresponding rank.

Rankings based on the input contribution to the reduction of the total deviance suggest four groups of inputs. The first group, including storage temperature and storage time at stage 3, represents the most important inputs with percentages of contribution to reduction in total deviance clearly higher than the other inputs. The second group represents inputs with medium importance, including storage time and lag period at stage 1. The group of inputs with low importance includes maximum density, generation time at stage 3, lag period at stage 3, and storage temperature at stage 1. There are five inputs in the group of inputs with no importance including inputs associated with the second stage of the growth process and generation time at stage 1.

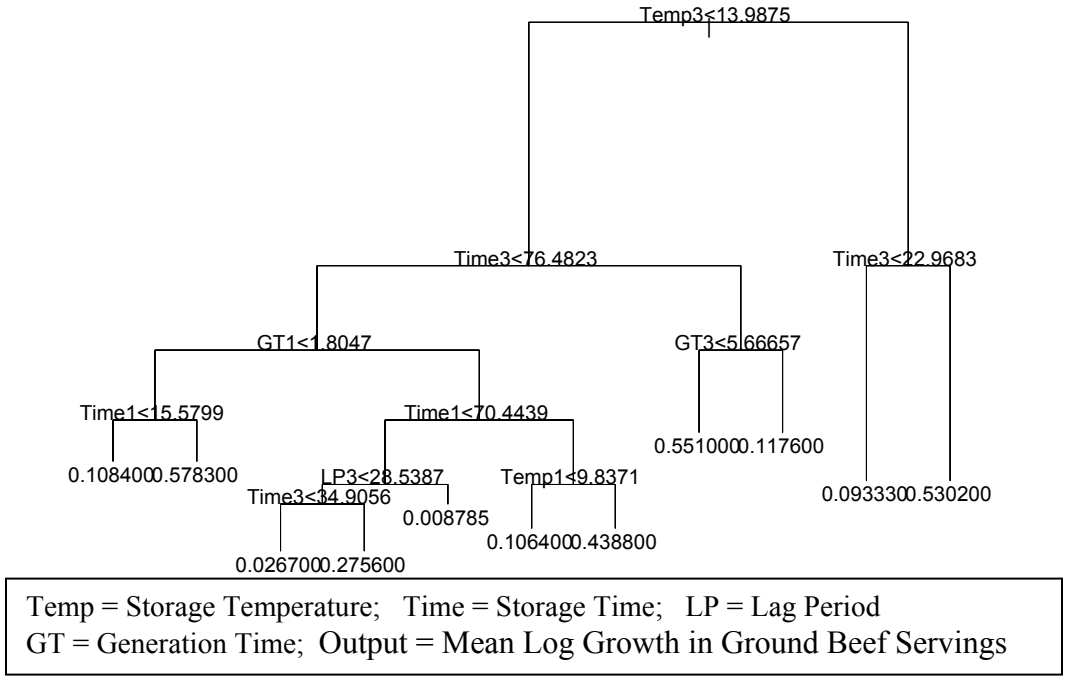



Figure 7-8. The Regression Tree for the Growth Estimation Part for One-Dimensional Variability and Uncertainty Analysis.

The rankings based on the contribution of inputs to the reduction in the total deviance are compared with those from visualization of the regression tree and complementary analyses in Section 7.3.1.3.

7.3.1.2 One-Dimensional Simulation of Variability and Uncertainty

This section presents the results of CART analysis applied to a one-dimensional probabilistic simulation in which variability and uncertainty are co-mingled, based upon the case study scenario described in Section 3.3.3. The results of CART analysis are depicted in Figure 7-8 in the form of a regression tree.

In CART analysis for this part the maximum number of leaves for the tree was specified as 11. Hence, in Figure 7-8 there are 11 mean responses presented in the regression tree. This number of nodes account for almost 85 percent of the variability in the output that can be captured if no restriction in the number of nodes was considered.

The results of the CART analysis for the co-mingled variability and uncertainty analysis are qualitatively similar to those for the variability only analysis of the previous section.

Specifically, the storage temperature at stage 3 is selected as the most important input. In this case a temperature of 14.0 $^{o}$C is the basis of the split, compared to 14.7 $^{o}$C in the previous case. The storage time at stage 3 is selected as the basis for the next split in both of the main branches of the tree. A high mean response of 0.53 is associated with the right most leaf based upon high values of both the temperature and storage time in stage 3, irrespective of the values of any other inputs. Another high mean response of 0.58 is associated with a stage 1 storage time of greater than 16 hours for a short stage 1 generation time and comparably low values of stage 3 storage time and temperature. Another large mean response of 0.55 is associated with large values of stage 3 storage time combined with the lower range of partitions with respect to stage 3 generation time. Thus, there are a variety of interactions among different inputs that can give rise to large growth. These interactions are influenced by nonlinearities and thresholds, as described in the previous section.

The CART results imply that the storage temperature at stage 3 is the most sensitive input, since it was selected as the basis for the first split in the tree. The storage time in stage 3 is deemed to be the second most important input, since it was selected in both of the second level nodes of the tree. Other inputs, such as the storage time and temperature at stage 1, the lag period in stage 3, and the generation time in stages 1 and 3, were selected in some of the lower nodes of the tree, especially on the left side of the tree. The left-most leaves correspond to comparably low values of stage 3 storage time and temperature. Thus, if these latter two variables have comparably low values, it is possible to have high growth depending on values of the other selected inputs.

In order to gain further insight regarding the sensitivity of the response to inputs other than the stage 3 storage time and temperature, regression analysis was conducted for two partitions of the data based upon a stage 3 storage temperature of 14 $^{o}$C. These results are summarized in Table 7-8.

For a stage 3 storage temperature of less than 14 $^{o}$C, the four most sensitive inputs are the stage 3 storage time, stage 1 storage time, stage 1 storage temperature, and stage 3 storage temperature. The confidence intervals of the first three inputs overlap. Thus, these inputs are of comparable importance. Other inputs for which there is a minor but statistically significant sensitivity include the generation time at stage 1 and lag periods at stages 3 and 1. Statistically insignificant inputs include the stage 2 storage time, lag period, and generation time.

Table 7-8. Regression Analysis Results for the Growth Estimation Part for One-Dimensional Variability and Uncertainty Analysis ($R^2$ = 0.48 for the first dataset and $R^2$ = 0.66 for the second dataset)

| Variable | Coefficient | 95% CI [(a)] | F Value | Pr>F | Rank |
|---|---|---|---|---|---|
| **Storage Temperature, Stage 3 <14$^{oC}$** | | | | | |
| Storage Temperature, Stage 1 | 0.38 | (0.37, 0.39) | 7,470 | <0.0001 | 3 |
| Storage Temperature, Stage 2 | 0.012 | (0.002,0.022) | 5 | 0.002 | 10 |
| Storage Temperature, Stage 3 | 0.25 | (0.24,0.26) | 1,935 | <0.0001 | 4 |
| Storage Time, Stage 1 | 0.386 | (0.38,0.39) | 17,620 | <0.0001 | 2 |
| Storage Time, Stage 2 | 0.004 | (-0.002,0.01) | 2 | 0.08 | --- |
| Storage Time, Stage 3 | 0.394 | (0.388,0.399) | 18,372 | <0.0001 | 1 |
| Maximum Density | 0.013 | (0.008,0.02) | 21 | <0.0001 | 9 |
| Lag Period, Stage 1 | -0.027 | (-0.035, -0.02) | 52 | <0.0001 | 7 |
| Lag Period, Stage 2 | -0.002 | (-0.011,0.006) | 0.3 | 0.6 | --- |
| Lag Period, Stage 3 | -0.032 | (-0.039,-0.024) | 63 | <0.0001 | 6 |
| Generation Time, Stage 1 | 0.077 | (0.068,0.085) | 302 | <0.0001 | 5 |
| Generation Time, Stage 2 | 0.006 | (-0.005,0.02) | 1 | 0.2 | --- |
| Generation Time, Stage 3 | 0.026 | (0.016,0.037) | 26 | <0.0001 | 8 |
| **Storage Temperature, Stage 3 >=14$^{oC}$** | | | | | |
| Storage Temperature, Stage 1 | 0.13 | (0.10,0.16) | 77 | <0.0001 | 3 |
| Storage Temperature, Stage 2 | -0.001 | (-0.04,0.03) | 0.2 | 0.7 | --- |
| Storage Temperature, Stage 3 | -0.119 | (-0.162,-0.076) | 30 | <0.0001 | 4 |
| Storage Time, Stage 1 | 0.15 | (0.13,0.17) | 240 | <0.0001 | 2 |
| Storage Time, Stage 2 | -0.002 | (-0.02,0.02) | 0.2 | 0.7 | --- |
| Storage Time, Stage 3 | 0.72 | (0.70,0.74) | 5,550 | <0.0001 | 1 |
| Maximum Density | 0.09 | (0.07,0.11) | 90 | <0.0001 | 7 |
| Lag Period, Stage 1 | -0.016 | (-0.04,0.009) | 2 | 0.08 | --- |
| Lag Period, Stage 2 | 0.006 | (-0.024,0.035) | 0.1 | 0.8 | --- |
| Lag Period, Stage 3 | -0.108 | (-0.142,-0.075) | 41 | <0.0001 | 5 |
| Generation Time, Stage 1 | 0.005 | (-0.025,0.034) | 0.1 | 0.8 | --- |
| Generation Time, Stage 2 | 0.008 | (-0.028,0.044) | 0.2 | 0.7 | --- |
| Generation Time, Stage 3 | 0.094 | (0.05,0.14) | 14 | 0.005 | 6 |

(a) CI = Confidence Interval for the coefficient

The results for the case of stage 3 storage temperature of greater than 14 $^{o}$C are qualitatively similar, although there are some quantitative differences. For example, in this case, the rankings assigned to the first and second most important inputs are considered to be substantially different, since the confidence intervals of the standardized regression coefficients for these two inputs do not overlap.

Taking into account both the CART and the regression results, a judgment is made that the top four inputs are as follows, in decreasing order of importance: (1) stage 3 storage temperature; (2) stage 3 storage time; (3) stage 1 storage temperature; and (4) stage 1 storage time. The first two are identified and ranked based upon the results of the regression tree. The latter two are implied by the results of the regression tree, but their importance was more clearly identified based upon the complementary regression analyses. Although some other inputs were identified as statistically significant in the regression analysis, their regression coefficients were substantially smaller than those of the storage time and temperature at stage 1. Therefore, they are judged to be of minor importance compared to the four inputs listed here.

Similar to the approach presented in Section 7.3.1.1, the amount of contribution of each input to the total reduction of the dataset deviance is considered as a sensitivity analysis index. The dataset provided for the commingled analysis of variability and uncertainty in the growth estimation part has a total deviance of 1008. If no condition is considered for the number of nodes in the tree, approximately 84 percent (i.e., 845.2) is addressed by the fitted regression tree. Nine inputs are selected in the regression tree. These inputs include storage times at stages 1 and 3, storage temperatures at stage 1 and 3, lap periods at stages 1 and 3, generation time at stages 1 and 3, and maximum density.

Table 7-9 summarizes the amount of reduction in the total deviance associated with selection of these inputs in the regression tree. This table indicates that there were 8 levels in the regression tree. There were multiple branches at a given level of the tree other than the first level. Hence, each input could appear several times in each level. Therefore, corresponding to each appearance of the input in the specific level of the tree, there is an associated numerical value representing the amount of reduction in the total deviance. The largest individual reduction in the total deviance is associated with the selection of storage temperature at stage 3 at the first splitting node. The total reduction in the deviance associated with each input is based upon the cumulative effect of repeated splits at lower levels of the regression tree.

Table 7-9. Reduction in Deviance Associated with Selected Inputs in the Regression Tree Generated in the Growth Estimation Part for the Variability Only Analysis

| Level of the Tree | Selected Inputs in the Regression Tree (1) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **$Time_1$** | **$Time_3$** | **$Temp_1$** | **$Temp_3$** | **MD** | **$LP_1$** | **$LP_3$** | **$GT_1$** | **$GT_3$** |
| **1st Level** | | | | 186.7 | | | | | |
| **2nd Level** | | 64.9<br>165.2 | | | | | | | |
| **3rd Level** | | 13.3 | | 17.6 | | | | 53.6 | 63.2 |
| **4th Level** | 31.9<br>49.6 | 12.0<br>19.4 | | | 6.1 | | | | 3.9<br>10.9 |
| **5th Level** | | | 33.5 | | | | 15.1<br>4.8 | | 3.6 |
| **6th Level** | 8.7 | 36.0 | | | | 7.6 | | 4.0 | |
| **7th Level** | 16.6 | 4.6<br>2.5 | 3.6 | | | 4.3 | | | |
| **8th Level** | 2.2 | | | | | | | | |
| **Sum** | 109 | 317.9 | 37.1 | 204.3 | 6.1 | 11.9 | 19.9 | 57.6 | 81.6 |
| **Percent of Contribution (2)** | 12.9 | 37.6 | 4.4 | 24.2 | 0.7 | 1.4 | 2.4 | 6.8 | 9.7 |
| **Rank** | 3 | 1 | 6 | 2 | 9 | 8 | 7 | 5 | 4 |

(1) Time = Storage Time, Temp = Storage Temperature, MD = Maximum Density, LP = Lag period, and GT = Generation Time. Subscript numbers indicate corresponding stage of the growth process.
(2) Total deviance of the dataset is 1008. The amount of deviance captured by the regression tree is 845.4.

The rankings in Table 7-9 are based on the amount of contribution of each input to the total reduction of the dataset deviance. The rankings indicate that storage time at stage 3 is the most important input, although this input is not selected at the first splitting node. Storage temperature at stage 3 is selected at the first splitting node and has a rank of second. The rankings indicate that there are four groups of inputs. The first group, corresponding to the most important inputs, includes storage time and storage temperature at stage 3. These inputs have clearly a higher contribution to the reduction in the total deviance. The second group includes three inputs with medium importance, including generation time at stages 1 and 3 and storage time at stage 1. The third group represents the inputs with low importance including maximum density, lag periods at stages 1 and 3, and storage temperature at stage 1. Finally there are four inputs that are not selected in the regression tree representing inputs with no importance. Inputs associated with the second stage of the growth process include in this group.

Table 7-10. Summary of the CART Analysis Results Based on Variability Only and Co-mingled Variability and Uncertainty Analyses

| Variable | Ranks [5] | | | |
|---|---|---|---|---|
| | Analysis 1[1] | | Analysis 3[2] | |
| | Visual | Deviance | Visual | Deviance |
| Storage Temperature, Stage 1 | 3[4] | 5 | 3[4] | 6 |
| Storage Temperature, Stage 2 | NS[3],[4] | NS[3] | NS[3],[4] | NS[3] |
| Storage Temperature, Stage 3 | 1 | 1 | 1 | 2 |
| Storage Time, Stage 1 | 4[4] | 3 | 4[4] | 3 |
| Storage Time, Stage 2 | NS[3],[4] | NS[3] | NS[3],[4] | NS[3] |
| Storage Time, Stage 3 | 2 | 2 | 2 | 1 |
| Maximum Density | 9 [4] | 8 | 9[4] | 9 |
| Lag Period, Stage 1 | 5 [4] | 4 | 8[4] | 8 |
| Lag Period, Stage 2 | NS[3],[4] | NS[3] | NS[3],[4] | NS[3] |
| Lag Period, Stage 3 | 8[4] | 6 | 7[4] | 7 |
| Generation Time, Stage 1 | 7[4] | NS[3] | 5[4] | 5 |
| Generation Time, Stage 2 | NS[3],[4] | NS[3] | NS[3],[4] | NS[3] |
| Generation Time, Stage 3 | 6[4] | 7 | 6[4] | 4 |

(1) Ranks based on the variability only analysis.
(2) Ranks based on the one-dimensional co-mingled variability and uncertainty analysis.
(3) NS = Not statistically significant.
(4) Ranked based upon the complementary regression analysis.
(5) Rankings are based on two sensitivity indices: (1) visual index with complementary analysis; and (2) deviance index

The rankings based on the contribution of inputs to the reduction in the total deviance are compared with those from visualization of the regression tree and complementary analyses in Section 7.3.1.3.

#### 7.3.1.3 Summary and Comparison of the Results of CART Analysis in the Growth Estimation Part

In this section, the results of the sensitivity analyses with CART applied to the growth estimation part based upon two different probabilistic analysis approaches are compared. The two approaches include simulation of only variability and simulation of both variability and uncertainty in a single dimension. These two approaches were described in the previous two sections. A summary of the results of the sensitivity analysis for each approach is given in Table 7-10. For each probabilistic approach results are presented based upon two sensitivity indices including visualization of the regression tree with using complementary analyses and the measure of sensitivity considering the amount of contribution of each input to the reduction of the total deviance.

The results from both approaches were comparable. The rankings for the top inputs were the same, and the same set of inputs was identified as statistically insignificant. Storage temperature and storage time at stage 3 were identified as the top two important inputs in both of the probabilistic approaches and with alternative sensitivity indices. Thus, the results in this instance are unambiguous regardless of which probabilistic simulation approach is employed. Moreover, two sensitivity indices presented in this section for ranking inputs approximately provide the same ranking with respect to the identification of the insignificant inputs and the most important inputs. Rankings based on visualization of the regression tree assume that the input selected in the first splitting node is the most important input. A case study provided in the growth estimation part for the one-dimensional co-mingled analysis of variability and uncertainty indicated that this assumption is not always valid. For example, it is possible that an input selected repeatedly in the lower levels of the tree could have a larger cumulative contribution to reduction in deviance compared to the input selected for the first split.

The sensitivity index based upon the contribution of each input to total reduction in deviance can give slightly different insights than a direct inspection of the tree. For example, although storage temperature at stage 3 appears at the first split in the tree, this input is not associated with the largest total reduction in deviance. The input that is the basis for the first split in the tree is typically associated with the largest single incremental reduction in deviance, but not necessarily the largest cumulative reduction in the deviance. Thus, the first variable selected in the tree is often an important input, but may not necessarily be the most important input in every case. It appears to be the case that the input selected for the first split in the tree will typically be an important input, even if it is not the most important input. Since these findings are based upon only two case studies, additional evaluations should be performed in order to recommend the best sensitivity index for CART. Moreover, rankings obtained based on other sensitivity analysis methods, such as regression analysis and ANOVA, can be compared with those provided based on CART. These comparisons are presented in Chapter 11.

### 7.3.2 Application of CART to the Cooking Effect Part

As explained in Section 3.4.3.2, inputs for the cooking effect part inputs include cooking temperature, precooking treatment, and cooking place. Distributions for these inputs are summarized in Table 3-13. The output in the cooking effect part is the mean log reduction in the number of *E. coli* organisms. For the cooking effect part there is a one-dimensional variability

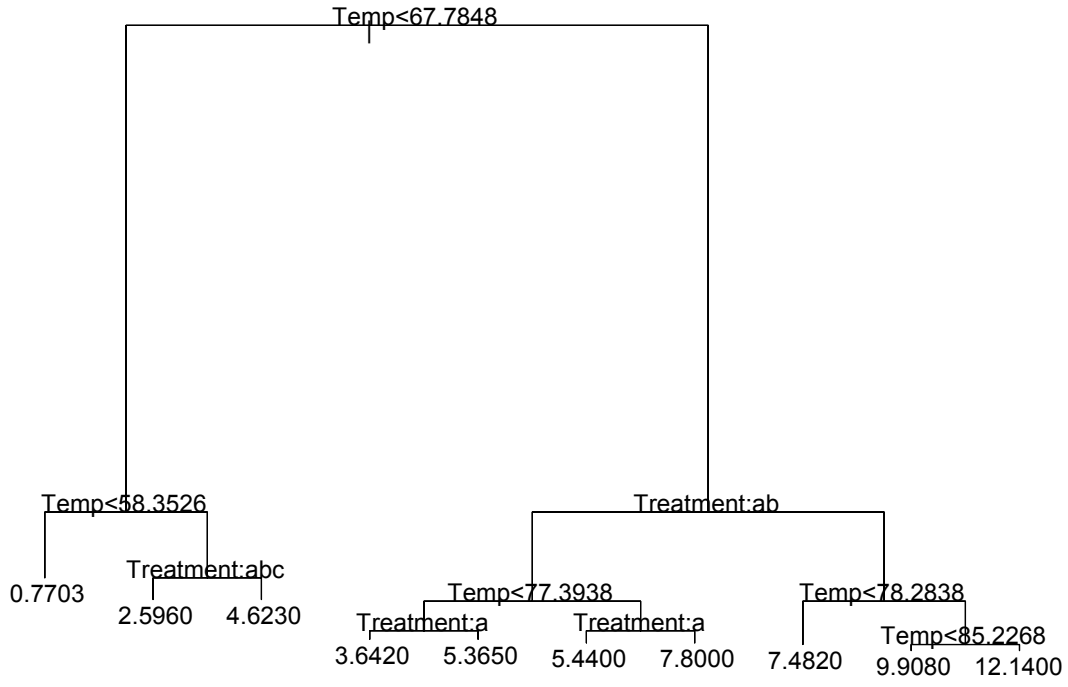



Figure 7-9. The Regression Tree for the Cooking Effect Part.

simulation with 65,000 iterations, as explained in Section 3.3.3. The results of CART analysis are depicted in Figure 7-9 in the form of a regression tree.

Figure 7-9 illustrates that the cooking temperature is the most important input. This is the first input selected. Furthermore, the proportional reduction in deviance is largest for the first partitioning of the database than for any of the subsequent partitions under the two main branches. The cooking temperature appears repeatedly throughout many of the lower nodes of the tree, which is also indicative of the importance of this input. The precooking treatment appears several times in the tree, suggesting that this input is of secondary importance. There are nine levels of precooking treatment. The partitions with respect to this input typically separate the first few such treatments from the remainder. For example, in the second level of nodes on the right side of the tree, which is conditional on a cooking temperature of greater than 68 $^{o}$C,

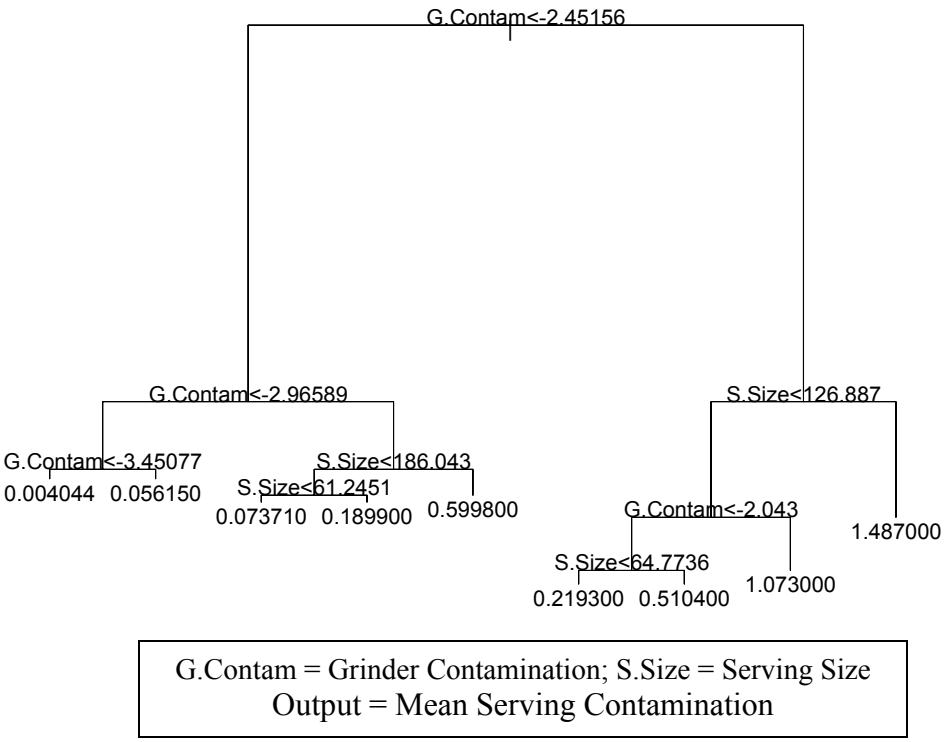



Figure 7-10. The Regression Tree for the Serving Contamination Part in Summer.

treatments a and b are partitioned into one data set and the remaining treatments *c* through *i* are partitioned into another data set.

The mean log reduction in *E. coli* organisms is greatest for the right-most leaf of the tree, corresponding to an average reduction of a factor of 12.14. This average reduction is based upon a cooking temperature of greater than 85 $^{o}$C and cooking pretreatments of *c* through *i*. In contrast, the lowest reduction is based upon a cooking temperature of less than 58 $^{o}$C irrespective of the type of pretreatment, as indicated by the left-most leaf of the tree. For cooking temperatures between 58 and 68 $^{o}$C, the average log reduction ranges from 2.60 to 4.62 depending upon the cooking pretreatment. Interactions between the cooking temperature and the pretreatment for cooking temperatures greater than 68 $^{o}$C are lead to average reductions ranging from 3.64 to 12.14. The larger average reductions are typically associated with higher cooking temperatures and with pretreatments other than *a* and *b*.

Overall, it is clear that the cooking temperature is the most important input but that there is also an important interaction between the cooking temperature and the cooking pretreatment.

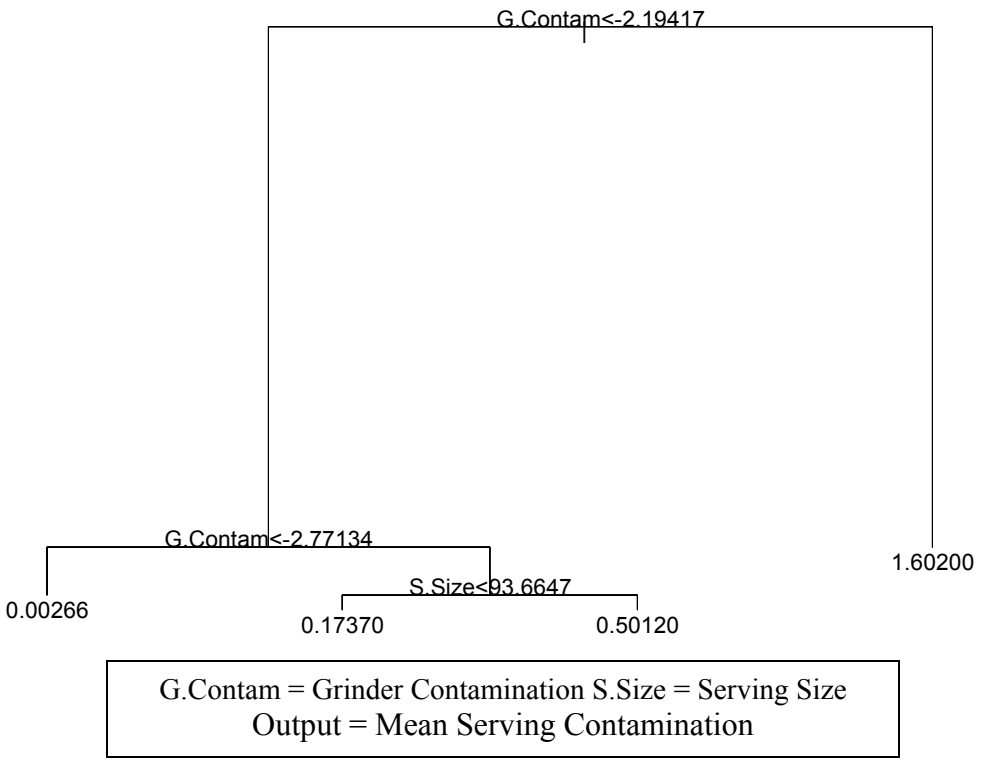



Figure 7-11. The Regression Tree for the Serving Contamination Part in Winter.

The partitioning of the data with respect to temperature suggests that there could be a temperature threshold below which there is little reduction in *E. coli* organisms because of cooking. In particular, for temperatures less than 58 $^{o}$C the log reduction in *E. coli* organisms was substantially smaller than for any other temperature range. The interactions between pretreatment and temperature for higher temperatures suggests that both of these inputs are of importance with respect to obtaining the highest possible reduction.

### 7.3.3 Application of CART to the Serving Contamination Part

As explained in section 3.4.3.1, inputs to the serving contamination part include the ground beef consumption type, serving size, eating location, consumer age, and grinder contamination. Distributions for these inputs are summarized in Table 3-12. The output in this part is the mean serving contamination prior to cooking. The case scenario in Section 3.3.3 focused on the serving contamination during high and low prevalence seasons. The results of CART analysis are depicted in Figures 7-10 and 7-11 for high and low prevalence seasons,

respectively, in the form of regression trees. No restriction was specified for the number of nodes in the regression tree. Hence, the mean responses presented in Figures 7-10 and 7-11 account for all of the variability in the output that could be captured by partitioning the dataset.

For the summer session, the first node in the regression tree subdivides the dataset into two divisions as follows: (1) grinder contamination less than –2.45 logs; and (2) grinder contamination greater than or equal to –2.45 logs. Data are further subdivided in lower levels based on values of the grinder contamination and serving size. The highest mean serving contamination of 1.5 as shown in the right-most leaf corresponds with high values of grinder contamination and large serving size.

The results in Figure 7-10 indicate that the mean serving contamination in summer is most sensitive to the grinder contamination, because this input is placed in the first node of the regression tree. Furthermore, there is large reduction in deviance for the first node as depicted by the long vertical distance of the first branches. Hence, the grinder contamination is ranked first. The serving size is considered as the second important input, because it is placed in lower nodes of the tree. The selection of just grinder contamination and serving size in the tree indicates that other inputs are not important based on the CART analysis.

For the winter session as shown in Figure 7-11, the first node in the regression tree subdivides the dataset into two divisions as follows: (1) grinder contamination less than –2.2 logs; and (2) grinder contamination greater than or equal to –2.2 logs. For the latter case, the mean serving contamination is 1.6 *E. coli* organisms. When the grinder contamination is less than –2.2 logs, the dataset is subdivided twice based on the grinder contamination and the serving size.

The results in Figure 7-11 indicate that the mean serving contamination in winter is most sensitive to the grinder contamination, because this input is placed in the first node of the regression tree. Furthermore, partitioning the data based upon this input leads to a large reduction in deviance, as illustrated by the comparably long vertical distance of the first level of branches compared to the lower levels of branches. Hence, the grinder contamination is ranked first. The serving size is considered as the second important input, because it is placed in lower nodes of the tree. Selection of just grinder contamination and serving size in the tree indicates that other inputs are not important based on the CART analysis.

The results for serving contamination were qualitatively similar for both the summer and winter seasons. In both cases, the grinder contamination is clear the most important input. The serving size is of secondary importance. Other inputs, such as eating location and consumer age, were not selected in the tree and therefore are comparably less important. Therefore, the findings regarding key sensitive inputs are robust with respect to season.

## 7.4 Evaluation of CART as a Sensitivity Analysis Method Based on Applications to the *E. coli* Model

In this chapter CART was applied to specific modules and parts of the *E. coli* model in order to identify the most important factors influencing the response of selected outputs. CART is a powerful method that is able to address both qualitative and quantitative inputs without any pre-processing of the dataset. Moreover, CART does not assume a specific functional relation between the model inputs and the model response. Hence, for models that have nonlinearity or thresholds application of CART does not force any under-estimation or over-estimation regarding the sensitivity of the output to each input.

CART does not have a specific sensitivity index. The ranking of the inputs in CART is typically based on visualization of the regression tree and judgment. For example, the regression trees indicate the proportional reduction in deviance associated with each node based upon the vertical distance of the branches. In some cases, the application of other sensitivity analysis methods as a complement to CART is needed to gain insight regarding the rank of each input. Because CART does not produce a sensitivity index similar to those of methods such as ANOVA or regression analysis, it is difficult to automate CART for application to many iterations. For example, when variability is simulated separately for multiple realizations of uncertainty in a two-dimensional probabilistic framework, it is difficult to summarize and compare the results of the CART analysis for each of the uncertainty realizations. Thus, the lack of a quantitative sensitivity index is a practical limitation that makes it difficult to automate CART for use with two-dimensional probabilistic analysis.

As an alternative sensitivity index, the contribution of each input to the reduction of total deviance of the dataset was presented in two case studies in the growth estimation part. This approach provided approximately the same ranking in comparison with the visualization of the regression tree incorporating the results of the complementary sensitivity analysis methods. In one case, the top ranked input as identified by inspection of the tree was fund to have the second

largest contribution to reduction in total deviance, but was also identified as substantially more important than the next ranked input. However, in order to further explore the validity of this sensitivity index more case studies should be performed to compare and evaluate rankings based on these approaches.

The software used for CART analysis does not directly provide the amount of contribution of each input to the reduction of the total deviance. Hence, the output provided by the software should be analyzed in order to estimate this sensitivity index. Evaluation of this sensitivity index based on the output file provided by the software is time consuming and tedious. Therefore, only two case studies were provided for application of this alternative sensitivity index. For further evaluation of this approach, a code should be developed to automate the process of ranking the inputs based on this alternative sensitivity index.

As noted in specific examples throughout this chapter, CART is able to respond in an intuitively appropriate manner to nonlinearities, thresholds, and interactions among inputs. For example, with respect to the cooking effect, it is clear from the CART analysis that low cooking temperatures do not provide a substantial reduction in *E. coli* organisms. However, for high temperatures, the reduction is substantial but depends also on the type of pretreatment used. Thus, this is an example in which the model responds in a nonlinear manner, has an apparent threshold, and has an important interaction between two inputs. This type of insight would be difficult to obtain with some of the other sensitivity analysis methods, such as linear standardized regression analysis.

Thus, although CART has limitations regarding development of a clear rank ordering of inputs, particularly for inputs other than the first or second most important ones, CART does provide critical insights regarding the combination of conditions that lead to either the highest or the lowest exposure (and, hence, risk). Opportunities to extend the utility of CART via development of a new sensitivity index should be explored.

# 8 APPLICATION OF SCATTER PLOTS TO THE *E. COLI O157:H7* MODEL

The objective of this chapter is to present the results of sensitivity analysis of the *E. coli* model based upon scatter plots. The details of the methodology for scatter plots analysis are provided in Section 2.3.1. Scatter plots are used to assess possible trends in the data and potentially complex dependencies between inputs and the outputs of interest. Scatter plots for different modules and parts of the *E. coli* model are provided for selected important inputs that were identified based on other sensitivity analysis methods.

This chapter contains four sections. Section 8-1 presents scatter plots for the production module. Section 8-2 presents scatter plots for the slaughter module and Section 8-3 presents the scatter plots for the preparation module. In Section 8-4, the method of using scatter plots for the sensitivity analysis is evaluated and the advantages, disadvantages and key criteria for application of this method are summarized.

## 8.1 Application of Scatter Plots to Production Module

In the production module scatter plots are provided for four parts, including the feedlot prevalence, within feedlot prevalence, breeding herd prevalence, and within breeding herd prevalence. The scatter plots are presented for each of these four parts in Sections 8.1.1 to 8.1.4, respectively.

### 8.1.1 Feedlot Prevalence Part

The feedlot prevalence part in the production module is explained in Section 3.2.1 and inputs for this part are given in Table 3-9. The output of interest is the median feedlot prevalence. There is a one-dimensional uncertainty simulation in this part as discussed in Section 3.3.1. Based upon the results of other sensitivity analysis methods, such as ANOVA and regression analysis, the study and herd sensitivity were identified as the two most important inputs. The study is a qualitative variable. Thus, for purposes of developing a scatter plot, the median feedlot prevalence was plotted versus herd sensitivity for each of the four studies. The result is shown in Figure 8-1. The number of data points in the figure is 10,000.

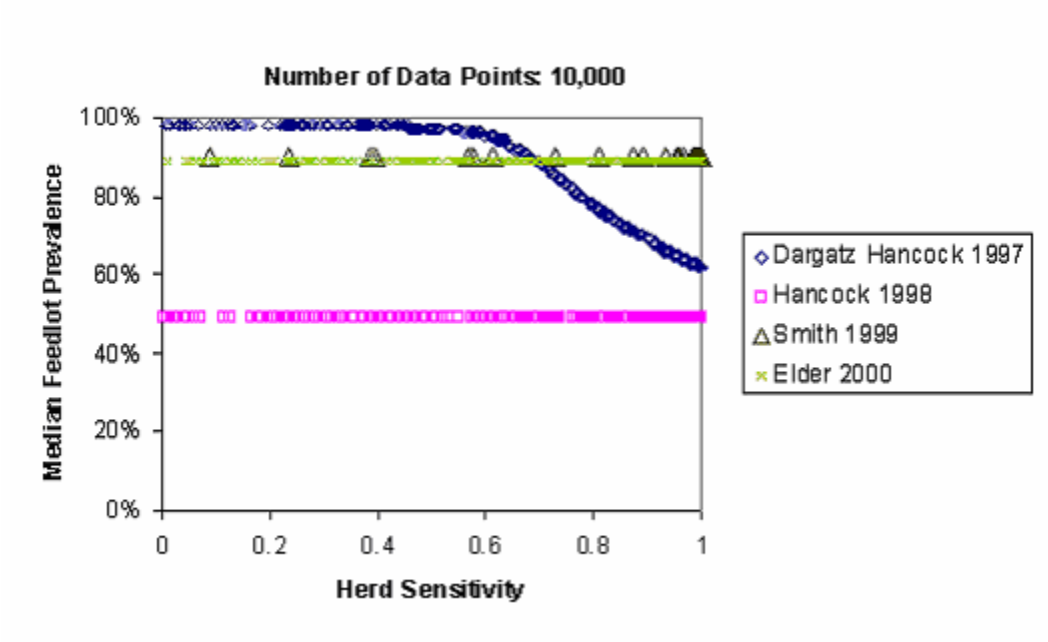



Figure 8-1. Scatter Plot for the Median Feedlot Prevalence Versus the Herd Sensitivity for Dargatz Hancock 1997, Hancock 1998, Smith 1999, and Elder 2000.

Figure 8-1 depicts that for three of the studies (Hancock 1998, Smith 1999, and Elder 2000) variation in the herd sensitivity does not have any effect on the median feedlot prevalence. However, for the Dargatz and Hancock (1997) study, an increase in the herd sensitivity leads to a decrease in the output above an apparent threshold. The definition of the herd sensitivity in Section 3.2.1 indicates that with an increase in the herd sensitivity the feedlot prevalence should decrease. However, the distribution considered for the herd sensitivity depends on the number of samples collected within herds and the detectable prevalence of infected animals in the infected herds. The median feedlot prevalence depends on the number of feedlot tested and number of positive feedlots. The information regarding each study in Table 3-1 indicates that the Dargatz and Hancock (1997) study has the highest number of tested cattle and lowest ratio of positive cattle to tested cattle among other studies. Moreover, the Dargatz and Hancock (1997) study has the highest number of tested feedlots. Thus, the study characteristics regarding the number of samples affected the way each study responds to the variation of the herd sensitivity.

The scatter plot in Figure 8-1 implies that the choice of study has a comparable impact on the median feedlot prevalence compared to the range of values for the herd sensitivity. For example, for a herd sensitivity of 0.5, the output varies from approximately 49 to 97 percent

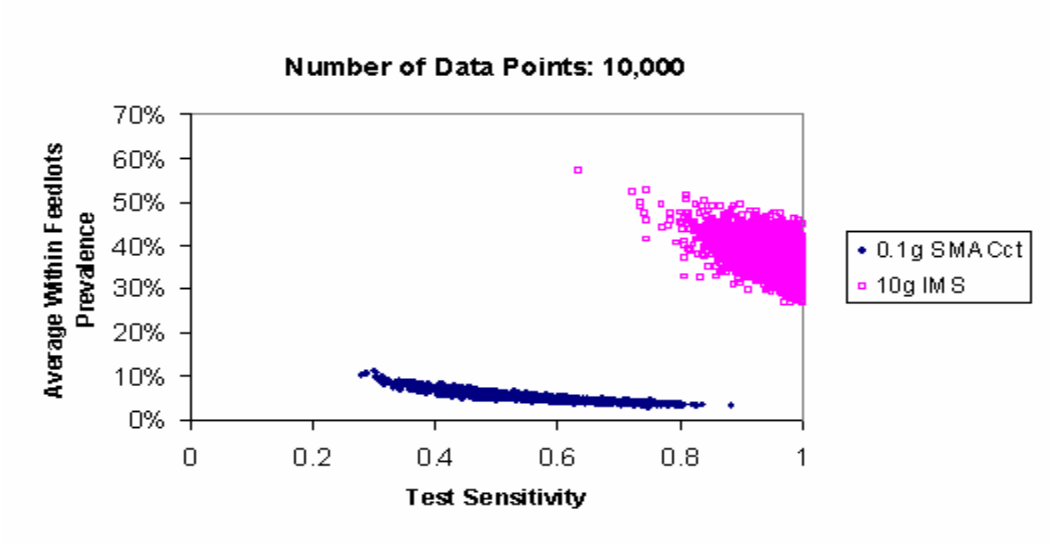



Figure 8-2. Scatter Plot for the Average Within Feedlots Prevalence Versus the Test Sensitivity for '0.1g SMACct' and '10g IMS' Testing Methods in Summer.

depending upon the choice of study, or a range of approximately 48 percentage points. In contrast, for a given choice of study, such as Dargatz and Hancock (1997), the output varies between approximately 60 and 100 percent, or a range of approximately 40 percentage points. Thus, the typical range of variation in the median feedlot prevalence is comparable with respect to values for the herd sensitivity as it is with respect to the choice of study. Moreover, there is an interaction effect between the study and the herd sensitivity, because for some studies the herd sensitivity does not have any effect, while for Dargatz and Hancock (1997) the response varies based on different values of the herd sensitivity.

### 8.1.2 Within Feedlot Prevalence Part

The within feedlot prevalence part is described in Section 3.2.1 and its inputs are summarized in Table 3-9. The output of interest is the average within feedlot prevalence in the high and low prevalence seasons. The case study for this part is based upon a one-dimensional simulation of uncertainty as discussed in Section 3.3.1. Based upon results with other sensitivity analysis methods, the apparent within feedlot prevalence and the test sensitivity were identified as the most important inputs and are the focus of analysis using scatter plots. Moreover, summer was identified as the season with higher average within feedlot prevalence. Thus, the average within feedlots prevalence for the summer season, which is the high prevalence season, is plotted

versus test sensitivity in Figure 8-2 and versus the apparent within feedlots prevalence in Figure 8-3. Each figure includes 10,000 randomly simulated data points.

Figure 8-2 indicates that the average within feedlot prevalence is relatively insensitive to test sensitivity for the "0.1 g SMA Cct" test method compared to the "10g IMS" method. In the latter case, the test sensitivity ranges between approximately 0.7 and 1.0, corresponding to output values ranging from 30 percent to more than 50 percent. The test sensitivity for the former method is generally lower, with a range from approximately 0.3 to 0.9, with a corresponding range of output values from as high as 10 percent to as low as three percent. For both methods, an increase in the test sensitivity is associated with a decrease in the average within feedlot prevalence.

The scatter plot in Figure 8-2 indicates that "10g IMS" testing method presents more accuracy in testing cattle in feedlots. Test sensitivity for "10g IMS" testing method ranges between almost 0.75 and 1.0, while for the "0.1g SMACct" testing method the test sensitivity varies between 0.25 and 0.9, indicating lower accuracy of the method.

The generated values for apparent within feedlots prevalence depend on the characteristics of the study such as number of samples within the infected feedlot. The scatter plot in Figure 8-3 indicates that there is an approximate linear relation between the average within feedlot prevalence and the apparent within feedlot prevalence. The parameters of the distribution for the apparent within feedlot prevalence given in Table 3-9 vary for each study. These parameters include the number of positive animals in a feedlot and the number of cattle tested in positive feedlots. Thus, the range of generated values for the apparent within feedlot prevalence varies for different study levels. This pattern is depicted in Figure 8-3 with discontinuity in the range of generated values for the apparent within feedlot prevalence and different slopes of the response in each range.

### 8.1.3 Breeding Herd Prevalence Part

The breeding herd prevalence part in the production module is explained in Section 3.2.1 and inputs for this part are given in Table 3-9. The output of interest is the median breeding herd prevalence. There is a one-dimensional uncertainty simulation in this part as discussed in Section 3.3.1. Based upon the results of other sensitivity analysis methods, such as ANOVA and regression analysis, study and herd sensitivity were identified as the two most important inputs. Thus, for purposes of developing a scatter plot, the median breeding herd prevalence was

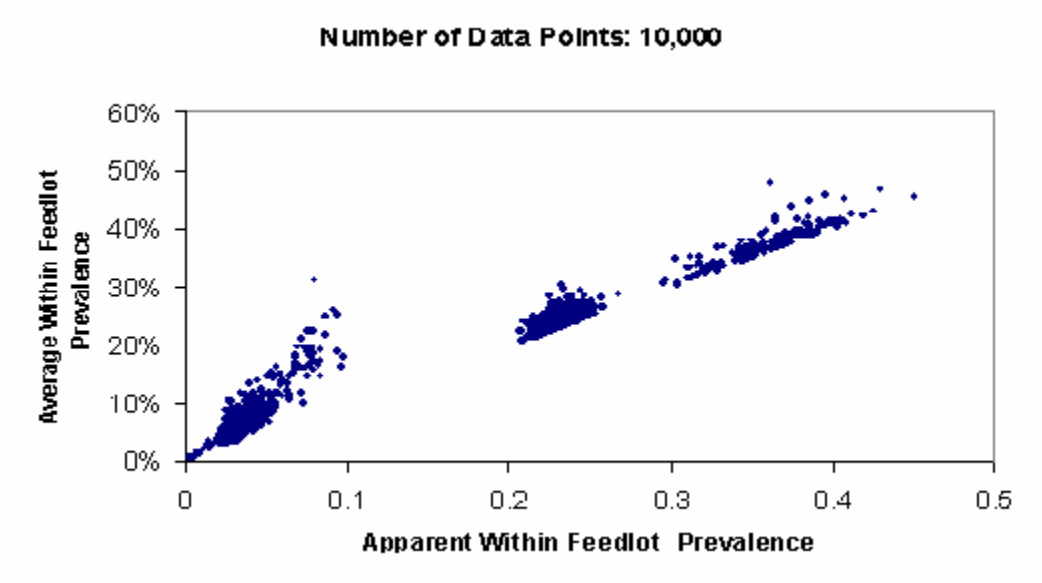



Figure 8-3. Scatter Plot of the Average Within Feedlot Prevalence Versus the Apparent Within Feedlot Prevalence in Summer.

plotted versus herd sensitivity for each of the seven studies. The result is shown in Figure 8-4. The number of data points in the figure is 10,000.

Figure 8-4 depicts that for four of the studies (Sargeant 2000, Hancock 1998, Hancock/CFSAN 2001, and Hancock 1997a) variation in the herd sensitivity does not have any effect on the median breeding herd prevalence. However, for the Garber 1998 study, an increase in the herd sensitivity leads to a decrease in the output. The definition of the herd sensitivity in Section 3.2.1 indicates that with increase in the herd sensitivity the breeding herd prevalence should decrease. However, the distribution considered for the herd sensitivity depends on the number of samples collected within herds and the detectable prevalence of infected animals in the infected herds and the median breeding herd prevalence depends on the number of herds tested and number of positive herds. The information regarding each study in Table 3-3 indicates that the Garber (1998) has the highest number of tested cattle and lowest ratio of positive cattle to tested cattle among other studies. Moreover, Garber (1998) has the highest number of tested herds. Thus, the study characteristics regarding the number of samples affected the way each study responds to the variation of the herd sensitivity.

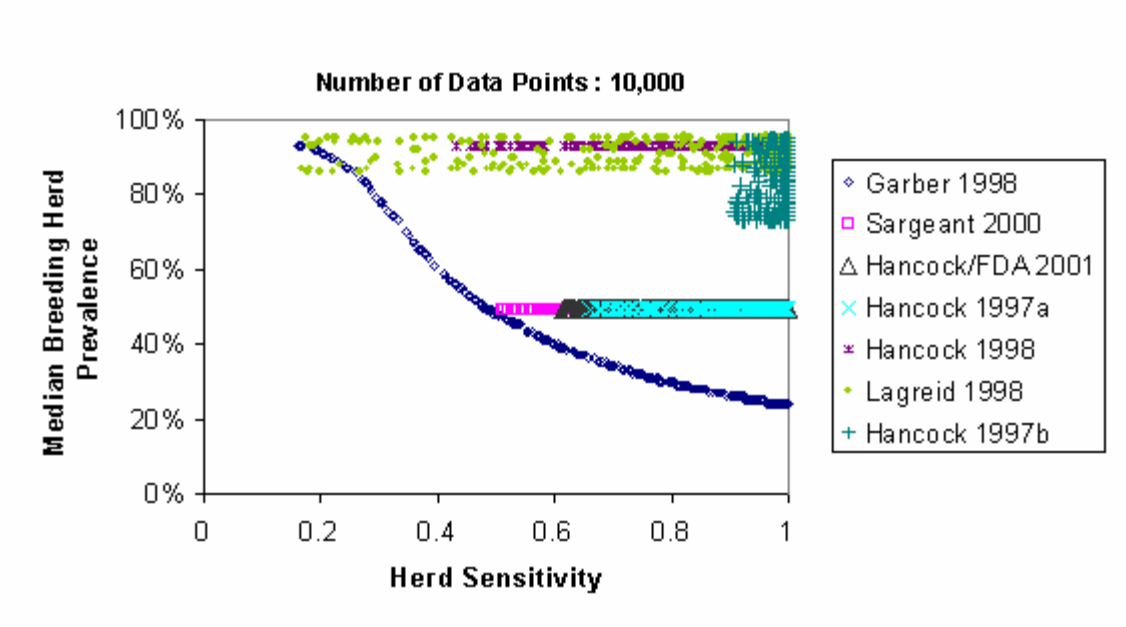



Figure 8-4. Scatter Plot for the Median Breeding Herd Prevalence Versus the Herd Sensitivity for Garber 1998, Sargeant 2000, Hancok/CFSAN 2001, Hancock 1997a, Hancock 1998, Lagreid 1998, and Hancock 1997b.

The scatter plot in Figure 8-4 implies that the choice of study has a comparable impact on the median breeding herd prevalence compared to the range of values for the herd sensitivity. For example, for a herd sensitivity of 0.6, the output varies from approximately 30 to 95 percent depending upon the choice of study, or a range of approximately 65 percentage points. In contrast, for a given choice of study, such as Garber (1998), the output varies between approximately 25 and 70 percent, or a range of approximately 65 percentage points. Moreover, there is an interaction effect between the study and the herd sensitivity, because for some studies the herd sensitivity does not have any effect, while for Garber (1998) the response varies based on different values of the herd sensitivity.

### 8.1.4 Within Breeding Herd Prevalence Part

The within breeding herd prevalence part is described in Section 3.2.1 and its inputs are summarized in Table 3-9. The output of interest is the average within breeding herd prevalence in the high and low prevalence seasons. The case study for this part is based upon a one-dimensional simulation of uncertainty as discussed in Section 3.3.1. Based upon results with other sensitivity analysis methods, the apparent within breeding herd prevalence and the test sensitivity were identified as the most important inputs and are the focus of analysis using scatter plots. Moreover, summer was identified as the season with higher average within breeding herd

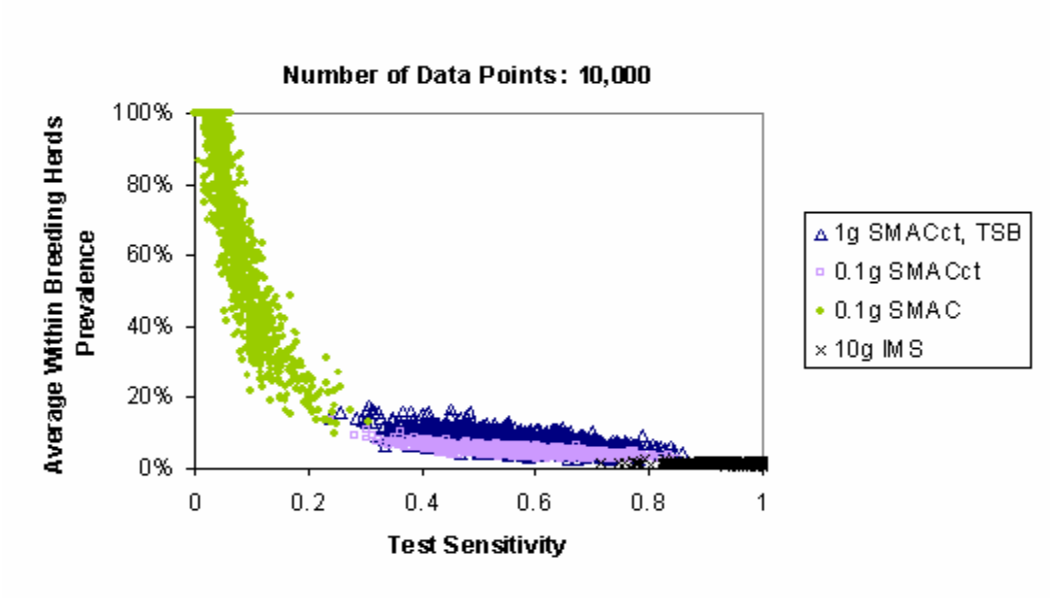



Figure 8-5. Scatter Plot for the Average Within Breeding Herd Prevalence Versus the Test Sensitivity for 1g SMACct TSB, 0.1g SMACct, 0.1g SMAC, and 10g IMS Testing Methods in Summer.

prevalence. Thus, the average within breeding herd prevalence for the summer season, which is the high prevalence season, is plotted versus test sensitivity in Figure 8-5 and versus the apparent within breeding herd prevalence in Figure 8-6. Each figure includes 10,000 randomly simulated data points.

Figure 8-5 indicates that the average within breeding herd prevalence is relatively insensitive to test sensitivity for the "0.1 g SMA Cct", "10g IMS", and "1g SMACct, TSB" test methods compared to the "0.1g SMAC" method. In the latter case, the test sensitivity ranges between approximately 0.05 and 0.025, corresponding to output values ranging from 10 percent to 100 percent. The test sensitivity for the former methods is generally higher, with a range from approximately 0.3 to 1.0, with a corresponding range of output values from as high as 18 percent to as low as three percent. For all methods, an increase in the test sensitivity is associated with a decrease in the average within breeding herd prevalence.

The scatter plot in Figure 8-5 indicates that the "10g IMS" testing method presents more accuracy in testing cattle in breeding herds. The test sensitivity for the "10g IMS" testing method ranges between 0.75 and 1.0, while for the "0.1 g SMA Cct", "0.1g SMAC", and "1g SMACct" testing methods the test sensitivity varies between 0.30 and 0.8, 0.05 and 0.25, and 0.30 and 0.8, respectively, indicating lower accuracy of these methods.

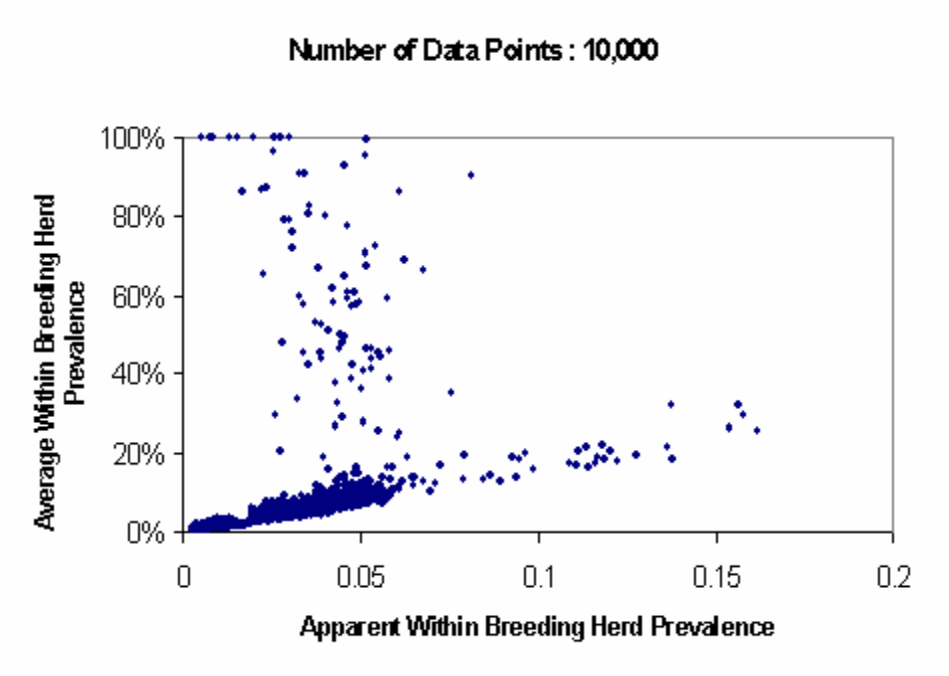



Figure 8-6. Scatter Plot for the Average Within Breeding Herds Prevalence Versus the Apparent Within Breeding Herds Prevalence in Summer.

The scatter plot in Figure 8-6 indicates that there is an approximate linear relationship between the average within breeding herd prevalence and the apparent within breeding herd prevalence. The parameters of the distribution for the apparent within breeding herd prevalence given in Table 3-9 vary for each study. These parameters include the number of positive animals in a herd and the number of cattle tested in positive herds. Thus, the range of generated values for the apparent within breeding herd prevalence varies for different study levels. This pattern is depicted in Figure 8-6.

## 8.2 Application of Scatter Plots to the Slaughter Module

Section 3.2.2 explains the slaughter module in the *E. coli* model. Inputs to the slaughter module are summarized in Table 3-10. The output of interest is the contamination in combo bins. The slaughter module includes both variability and uncertainty simulations. For scatter plots in the slaughter module a simulation of 650 variability and 100 uncertainty iterations was performed as a case representing all the possible values of inputs including both variability and uncertainty. The case study scenario for the slaughter module is focused upon steers and heifers in the high prevalence season.

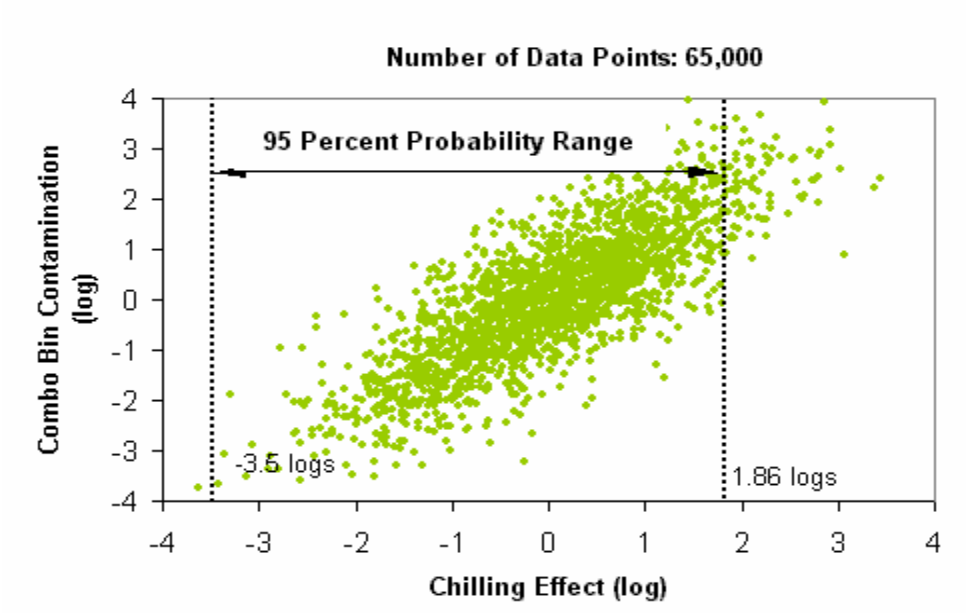


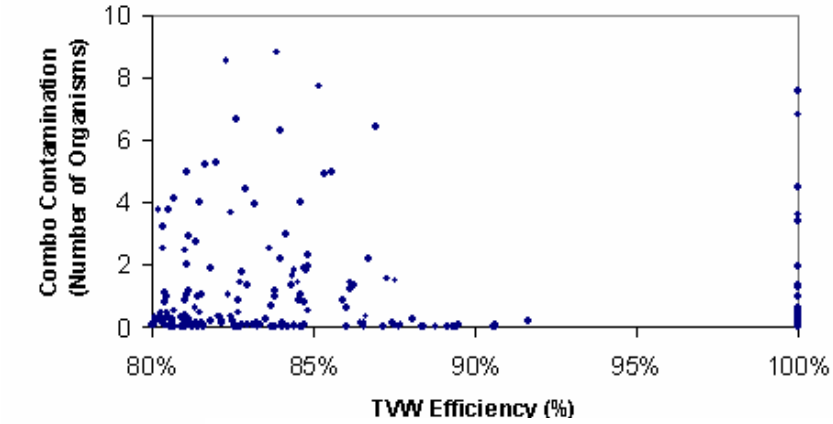

Figure 8-7. Scatter Plot for the Combo Bin Contamination from Steers and Heifers Versus the Chilling Effect in Summer.

Based upon results with other sensitivity analysis methods, chilling effect, number of organisms, and Trim/Vacuum/Wash efficiency were identified as the most important inputs and are the focus of analysis using scatter plots. In Figures 8-7 to 8-9 the scatter plots for chilling effect, Trim/Vacuum/Wash efficiency, and number of organisms are depicted, respectively.

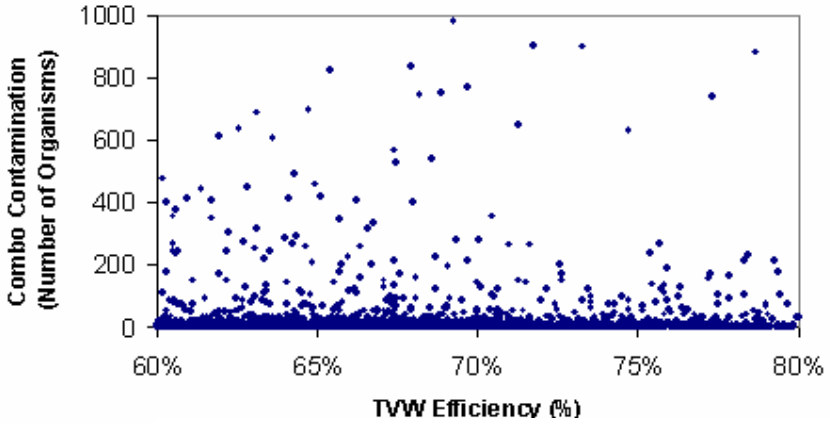
Figure 8-7 presents a scatter plot for combo bin contamination versus the chilling effect. Values of the chilling effect between –4 logs and 4 logs are depicted in this figure. There are a few data points with chilling effect of –5 logs representing cases with no contamination with *E. coli* organisms. These points are not depicted in Figure 8-7. Based on the temperature during the chilling process the number of *E. coli* organisms on carcasses might increase or decrease (FSIS, 2001). The scatter plot in Figure 8-7 depicts that there is a linear relationship between the combo bin contamination and the chilling effect. Not all values from the simulation are shown. The chilling effect ranges between –4 logs and 3 logs. The chilling effect, the combo bin contamination varies between –3 logs and 4 logs. The 95 percent probability range of the chilling effect is between –3.5 logs and 1.86 logs with a median value of 0.7 logs. The median combo bin contamination is –1.2 logs.

Figure 8-8 presents the scatter plots for the combo bin contamination versus the Trim/Vacuum/Wash efficiency. For this scatter plot, the values in the *X* axis representing the Trim/Vacuum/Wash efficiency are subdivided into three ranges: (1) TVW efficiency between 80% and 100%; (2) TVW efficiency 60% and 80%; and (3) TVW efficiency less than 60%. Classifying the generated values for the TVW efficiency facilitates the inference from the scatter plots. When there is a high efficiency in the decontamination step for contaminated carcasses using Trim/Vacuum/wash process (i.e., efficiency more than 80 percent), number of *E. coli* organisms in combo bins made from these carcasses is usually less than 2 *E. coli* organisms. There are also combo bins with near 1 log of contamination when using high efficiency in the decontamination step. With lower decontamination efficiency less than 80%, the number of *E. coli* organisms increases and most of the combo bins have contamination as high as 2.3 logs and as low as 1 log. There are also combo bins with approximately 3 logs of contamination. The scatter plot in Figure 8-8 implies that in order to keep the combo bin contamination less than 1 log, the efficiency of the decontamination step using the Trim/Vacuum/wash process should be above 80 percent.

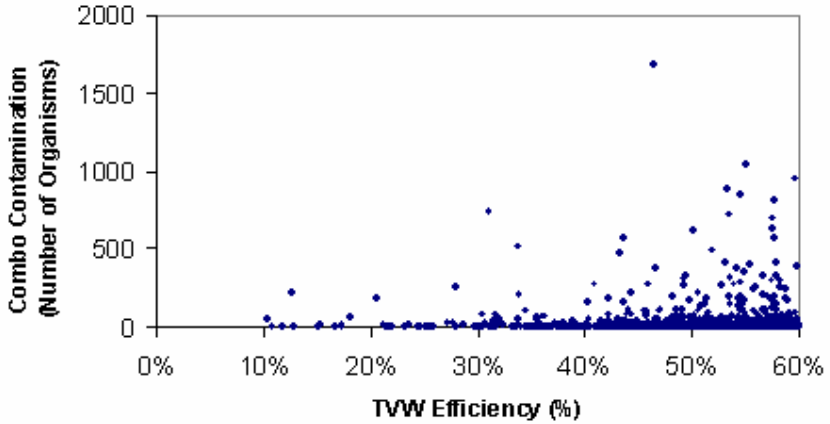
Figure 8-9 presents a scatter plot for combo bin contamination versus the number of organisms on contaminated carcasses. The scatter plot in Figure 8-9 depicts that there is a linear relation between the combo bin contamination and the number of organisms. The 95 percent probability for the range of the number of organisms is between –2 logs and 3.1 logs with the median value of 1.1 logs for the chilling effect. The median combo bin contamination is 0 log. Values of the number of organisms between –2 logs and 5 logs are depicted in this figure. There are a few data points with chilling effect of –5 logs representing cases with no contamination with *E. coli* organisms. These points are not depicted in Figure 8-9.




(a) TVW Efficiency Between 80% and 100%.




(b) TVW Efficiency Between 60% and 80%.




(c) TVW Efficiency Less Than 60%.

Figure 8-8. Scatter Plot for the Combo Bin Contamination for Steers and Heifers Versus the TVW Efficiency Effect in Summer.

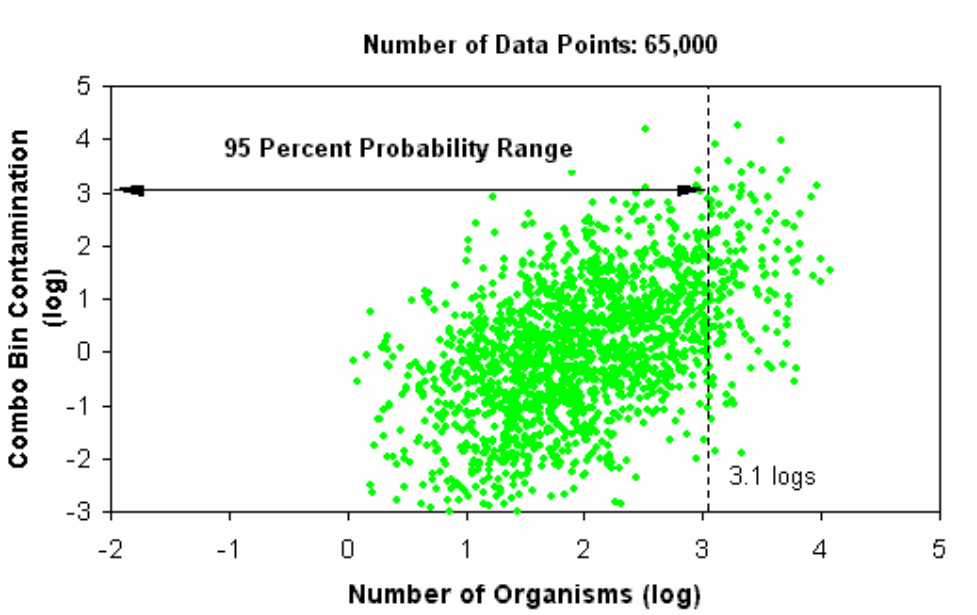



Figure 8-9. Scatter Plot for the Combo Bin Contamination from Steers and Heifers Versus the Number of Organisms on Contaminated Carcasses in Summer.

## 8.3 Application of Scatter Plots to the Preparation Module

In the preparation module scatter plots are prepared for three parts, including the growth estimation, the cooking effect, and the serving contamination parts. The scatter plots are presented for each of these three parts in Sections 8.3.1 to 8.3.3, respectively.

### 8.3.1 Application of Scatter Plots to the Growth Estimation Part

The growth estimation part in the preparation module is explained in Section 3.2.3 and inputs for this part are given in Table 3-11. The output of interest is the mean growth of the *E. coli* organisms in ground beef servings. The growth estimation part includes both variability and uncertainty simulations. For scatter plots in this part a simulation of 650 variability and 100 uncertainty iterations was performed as a case representing all the possible values of inputs including both variability and uncertainty. The generated values were co-mingled in one dimension for the analysis using the scatter plot.

Based upon results with other sensitivity analysis methods, storage time and storage temperature at stage 3 (i.e., home) were identified as the most important inputs and are the focus of analysis using scatter plots. In Figures 8-10 to 8-11 the scatter plots for the mean growth versus the storage time and the storage temperature at stage 3 are presented, respectively.

Figure 8-10 implies that when the storage time at home is more than 112 hrs there is no serving with zero growth in the number of *E. coli* organisms. In contrast, when the storage time is less than 112 hrs there are servings with no growth. This trend indicates that with the storage time less than 112 hrs it is possible to stop the growth of *E. coli* organisms in ground beef servings by controlling other inputs such as the storage temperature. For longer storage times no condition can stop the *E. coli* organisms from growing in ground beef servings. Although there is no clearly defined threshold, the scatter plot implies that growth is typically larger for the larger storage times, especially above 112 hrs.

Figure 8-11 depicts that with increase in the storage temperature the contamination in ground beef servings increases. For temperatures less than $10^{o}$C most of the servings have growth estimation of less than 0.1 logs. In contrast, when the storage temperature increases to a value between $10^{o}$C and $15^{o}$C, there are more ground beef servings with contamination above 0.5 logs due to the growth of *E. coli* organisms. Moreover, there is a gap in Figure 8-11, and it seems that the storage temperatures between $16^{o}$C and $18^{o}$C were not generated in the random simulation. This pattern happened because of the cumulative distribution considered for the storage temperature at home in the original *E. coli* model. This distribution could not generate values between $16^{o}$C and $18^{o}$C.

### 8.3.2 Application of Scatter Plots to the Cooking Effect Part

The cooking effect part in the preparation module is explained in Section 3.2.3 and inputs for this part are given in Table 3-13. The output of interest is the log reduction in the number of *E. coli* organisms due to cooking. There is a one-dimensional variability simulation in this part as discussed in Section 3.3.3. Based upon the results of other sensitivity analysis methods, such as ANOVA, the cooking temperature identified as the most important input. Thus, for purposes of developing a scatter plot, the log reduction in the number of *E. coli* organisms due to cooking was plotted versus the cooking temperature. The result is shown in Figure 8-12.

Scatter plot in Figure 8-12 implies that there is a linear relationship between the log reduction in the number of *E. coli* organisms due to cooking and the cooking temperature. The scatter plot presents several lines for the relationship. Each line represents a specific precooking treatment. The lines presented in the scatter plot are not parallel indicating that there is an interaction between the cooking temperature and the precooking treatment identified by the scatter plot. Because of the interaction, the response of the model differs for low and high cooking temperature depending upon choice of the precooking treatment.

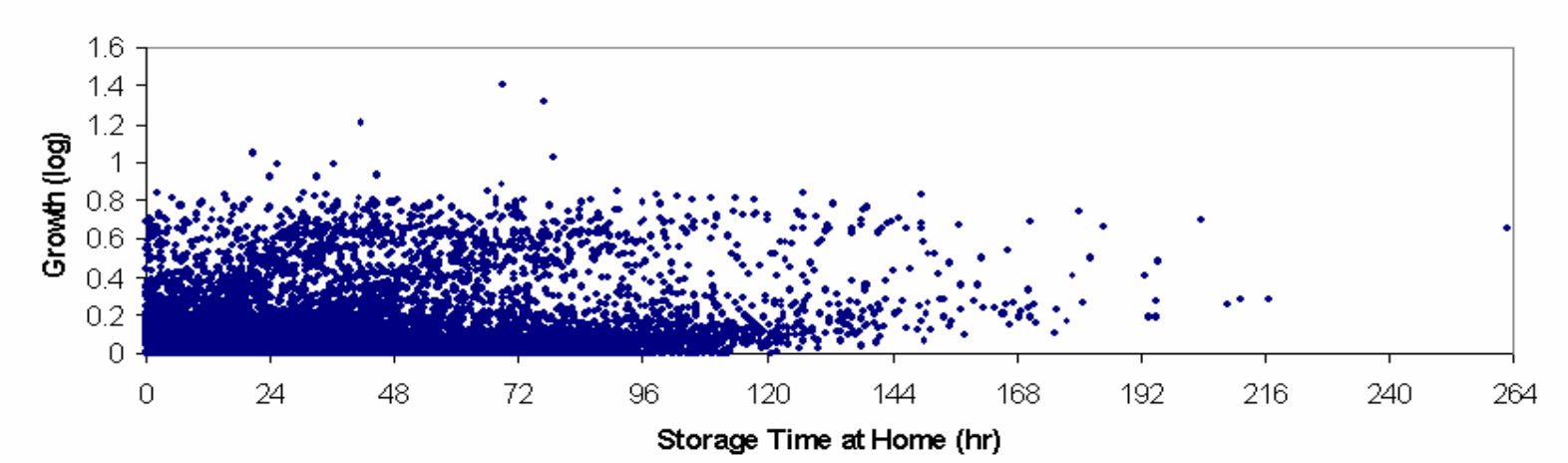



Figure 8-10. Scatter Plot for the Growth of *E. coli* Organisms versus the Storage Time at Home.

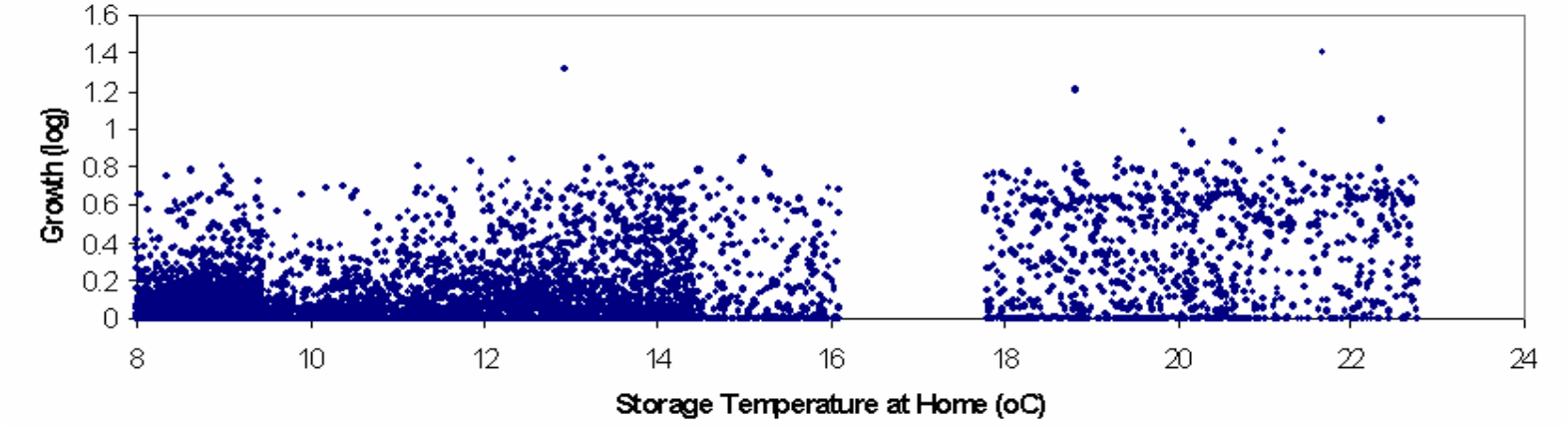



Figure 8-11. Scatter Plot for the Growth of *E. coli* Organisms versus the Storage Temperature at Home.

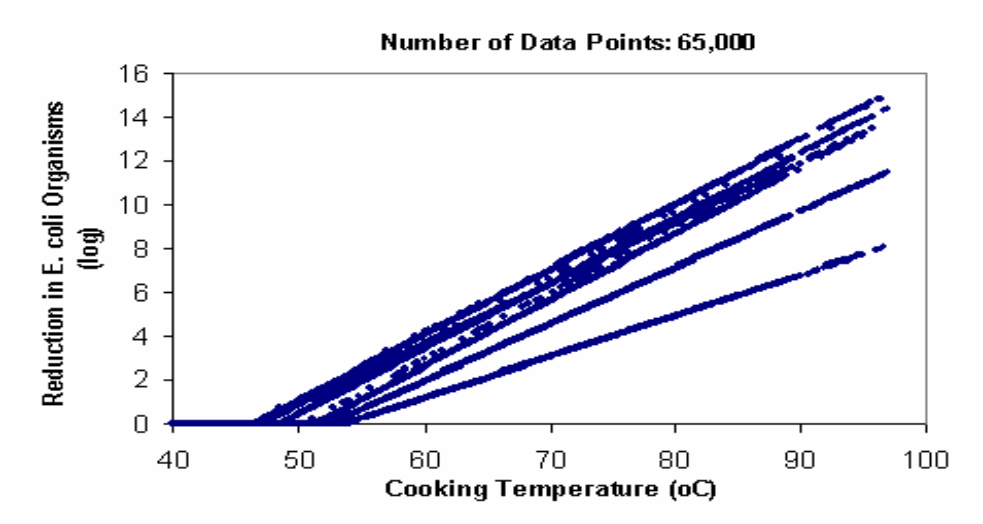



Figure 8-12. Scatter Plot for the Log Reduction in the Number of *E. coli* Organisms versus the Cooking Temperature at Home

For example, for a cooking temperature of 60°C the log reduction in the number of *E. coli* organisms due to cooking varies between 1 and 4 logs depending upon the choice of precooking treatment, or a range of 3 logs. In contrast, for a cooking temperature of 90 °C the log reduction in the number of *E. coli* organisms due to cooking varies between 7 and 13 logs depending upon the choice of precooking treatment, or a range of 6 logs.

There is a threshold in the response of the model to the cooking temperature. Cooking temperatures of less than a range between 47°C and 53°C, depending on the precooking treatment, have no effect on the reduction in the number of *E. coli* organisms.

### 8.3.3 Application of Scatter Plots to the Serving Contamination Part

The serving contamination part in the preparation module is explained in Section 3.2.3 and inputs for this part are given in Table 3-12. The output of interest is the mean serving contamination. There is a one-dimensional variability simulation in this part as discussed in Section 3.3.3. Based upon the results of other sensitivity analysis methods, such as regression analysis and CART, the grinder contamination was identified as the most important input. Thus, for purposes of developing a scatter plot, the mean serving contamination was plotted versus the grinder contamination in summer. The result is shown in Figure 8-13.

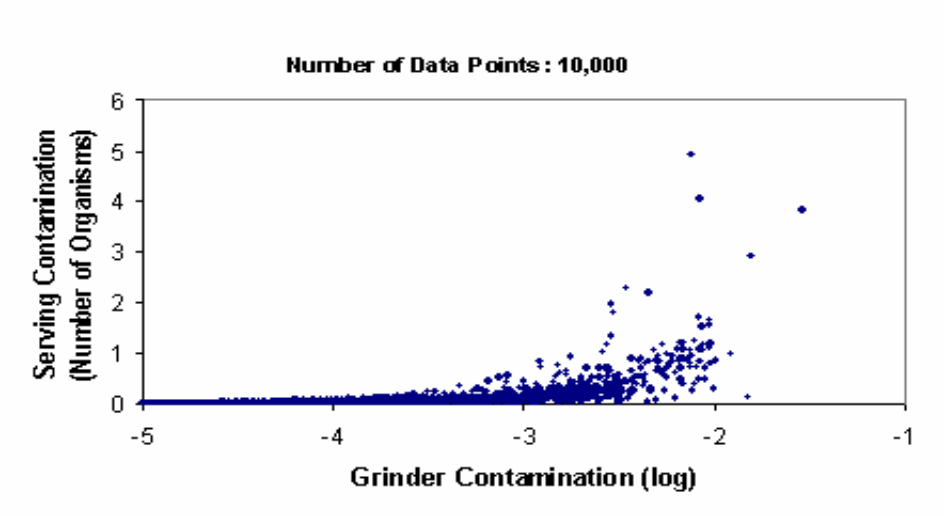



Figure 8-13. Scatter Plot for the Serving Contamination Versus the Grinder Contamination in Summer.

The scatter plot in Figure 8-13 implies 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. 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.

## 8.4 Evaluation of Scatter Plots as a Sensitivity Analysis Method Based on Applications to the *E. coli* Model

In Sections 8-1 to 8-3 scatter plots was applied to different modules and parts of the *E. coli* model. Scatter plots were implemented in order to clarify the relationship between the output and inputs such as non-linearity, thresholds, discontinuity, and interaction effects between inputs. Scatter plots cannot be used to explicitly rank the inputs. However, the possibility of clarifying special relationships is an advantage of this method of sensitivity analysis.

Non-linearity in the model response to specific inputs can be identified using scatter plots. For example, in Section 8.3.3 there is a non-linear response of the model to the variation of the grinder contamination. This trend cannot easily be identified by a sensitivity analysis method such as linear regression analysis, which assumes a specific functional relationship between output and inputs. Thus, identification of the non-linearity in the model response with scatter plots implies that the application of sensitivity analysis methods with pre-defined functional

relationships, such as regression analysis, should be accompanied with concern. It is possible that the results of the sensitivity analysis based on such methods are not reliable.

Thresholds in the model response to a specific input can be identified using scatter plots. For example, in Section 8.3.3 a threshold in the model response to the grinder contamination was identified. The ability to identify thresholds with scatter plots can be used in order to validate the results from other sensitivity analysis methods. For instance, Chapter 7 explained that CART could be used in order to identify thresholds, but that support from other sensitivity analysis methods are needed to justify the identified thresholds. For example, in Section 7.3.3 the grinder contamination of –2.45 logs was selected in the first node of the regression tree. Figure 7-10 indicated that the highest serving contamination was associated with cases that the grinder contamination was higher than –2.45 logs. However, based solely upon the CART analysis there was ambiguity as to whether this value could be considered as a threshold. Because approximately the same grinder contamination is identified as a threshold using the scatter plot, the result from CART is verified.

Scatter plots can be implemented in order to clarify interaction effects between inputs. For example, in Section 8.3.2 a scatter plot was used to identify the interaction between the cooking temperature and the precooking treatment in the cooking effect part.

It can be difficult to discern the interaction effect between inputs using scatter plots. In order to reveal such effects using scatter plots, it is helpful if one of the inputs is qualitative. In an example presented in Section 8.3.2, the precooking treatment was a qualitative input with 9 levels. Thus, 9 distinct patterns of points were identified representing precooking treatments. These distinct patterns facilitated the simultaneous evaluation of the precooking treatment and the cooking temperature effects on the log reduction in the number of *E. coli* organisms.

The capability of revealing interaction effects using scatter plots can be verified using other sensitivity analysis techniques such as ANOVA or regression that directly address the interaction effects between inputs. Table 5-26 indicates that there is a statistically significant interaction effect between the cooking temperature and the precooking treatment identified using ANOVA. This interaction effect was also identified using a scatter plot in Figure 8.12.

# 9 CONDITIONAL SENSITIVITY ANALYSIS FOR THE *E. COLI O157:H7* MODEL

The objective of this chapter is to present the results of sensitivity analysis of the *E. coli* model based upon the conditional sensitivity analysis method. The details of the methodology for conditional sensitivity analysis are provided in Section 2.3.2. Conditional sensitivity analysis is used to assess possible trends in the data and potentially complex dependencies between inputs and the outputs of interest. This method is applied to different modules and parts of the *E. coli* model for selected important inputs that were identified based on other sensitivity analysis methods.

This chapter contains four sections. Section 9-1 presents results of the conditional sensitivity analysis for the production module. Sections 9-2 and 9-3 present results of the conditional sensitivity analysis for the slaughter and preparation modules, respectively. In Section 9-4, the method of using conditional sensitivity analysis is evaluated and the advantages, disadvantages and key criteria for application of this method are summarized.

## 9.1 Application of the Conditional Sensitivity Analysis to the Production Module

In the production module conditional sensitivity graphs are provided for two parts, including the feedlot prevalence and within feedlot prevalence. These two parts were identified to have higher infection prevalence based on other sensitivity analysis methods. The conditional sensitivity graphs are presented for each of these two parts in Sections 9.1.1 and 9.1.2, respectively.

### 9.1.1 Application of Conditional Sensitivity Analysis to the Feedlot Prevalence Part

The feedlot prevalence part in the production module is explained in Section 3.2.1 and inputs for this part are given in Table 3-9. The output of interest is the median feedlot prevalence. There is a one-dimensional uncertainty simulation in this part as discussed in Section 3.3.1. Based upon the results of other sensitivity analysis methods, such as ANOVA and regression analysis, the study and herd sensitivity were identified as the two most important inputs. Thus, for purposes of developing conditional sensitivity plots, the median feedlot prevalence was plotted versus herd sensitivity for each of the four studies considering all other inputs conditioned at minimum, mean, and maximum values. The nominal values for each input were derived based on the input distribution. The nominal values for inputs to this part are given in Table 9-1.

Table 9-1. Nominal Values for Apparent Prevalence and Herd Sensitivity in the Feedlot Prevalence Part

| Variable | Study | Minimum | Mean | Maximum | Unit |
|---|---|---|---|---|---|
| Apparent Prevalence | Dargatz, Hancock 1997 | 0 | 2.8 | 100 | Percent |
| | Hancock 1998 | 0 | 3.7 | 100 | Percent |
| | Smith 1999 | 0 | 23 | 100 | Percent |
| | Elder 2000 | 0 | 36 | 100 | Percent |
| Herd Sensitivity | Dargatz, Hancock 1997 | 0 | 2.7 | 100 | Percent |
| | Hancock 1998 | 0 | 3.6 | 100 | Percent |
| | Smith 1999 | 0 | 22 | 100 | Percent |
| | Elder 2000 | 0 | 30 | 100 | Percent |

Although there are two specific distributions for herd sensitivity and apparent prevalence given in Table 3-9, the parameters of these distributions are function of the study features such as number of samples in feedlots. Hence, nominal values for these inputs differ for each study.

For the conditional sensitivity analysis, for each study three simulations with 2,000 iterations were performed. In each simulation, the herd sensitivity for that study was varied based on its distribution, while all other inputs were conditioned at either minimum, mean or maximum values.

Figure 9-1 depicts the conditional relationship between the median feedlot prevalence and the herd sensitivity for the Dargatz and Hancock (1997) study. The graph implies that a herd sensitivity less than approximately 0.5 does not have a substantial effect on the median feedlot prevalence. For cases in which the herd sensitivity is less than 0.5 the median feedlot prevalence varies in a narrow range of 99 to 100 percent. However, an increase in the herd sensitivity above 0.5 leads to a decrease in the output. Furthermore, there appears to be a large change in the slope of the curve near a herd sensitivity of 0.6. Thus, the value of 0.6 for the herd sensitivity can be considered as a threshold. Moreover, Figure 9-1 shows that the results are the same regardless of the values of other model inputs.

Figure 9-2 depicts the conditional relationship between the median feedlot prevalence and the herd sensitivity for the Hancock (1998) study. The graph implies that variation of the herd sensitivity does not have any effect on the median feedlot prevalence for this study. Moreover, The median feedlot prevalence varies between approximately 65% and 100%, or a

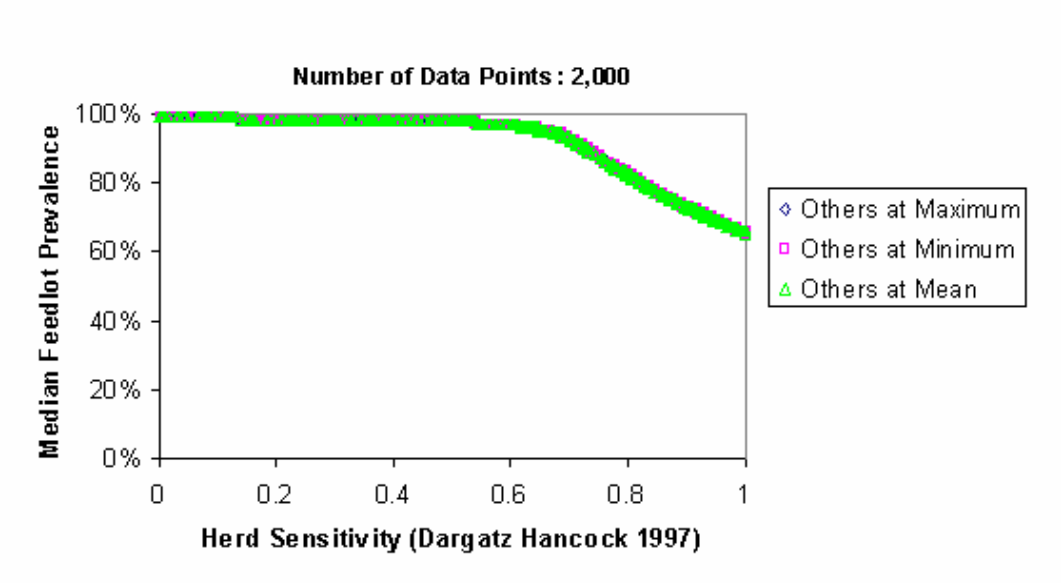



Figure 9-1. Conditional Sensitivity Analysis of the Herd Sensitivity, Dargatz, Hancock 1997 Study.

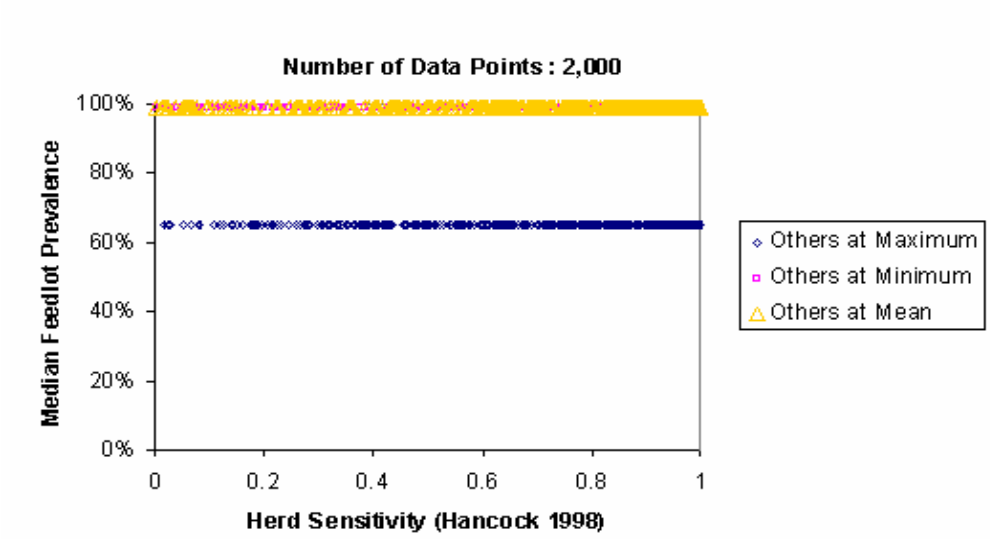



Figure 9-2. Conditional Sensitivity Analysis of the Herd Sensitivity, Hancock 1998 Study.

range of approximately 35 percentage points, depending on whether all other inputs are at the minimum or maximum values.

Figure 9-3 depicts the conditional relationship between the median feedlot prevalence and the herd sensitivity for the Smith (1999) study. For this study the variation of the herd sensitivity does not have any effect on the median feedlot prevalence. The median feedlot

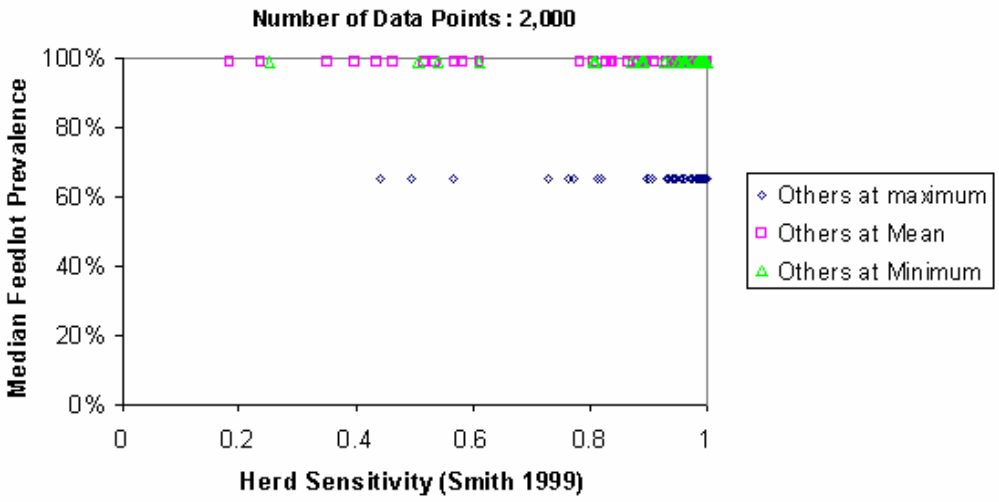



Figure 9-3. Conditional Sensitivity Analysis of the Herd Sensitivity, Smith 1999 Study.

prevalence varies between approximately 65% and 100%, or a range of approximately 35 percentage points, depending on whether all other inputs are at their maximum or minimum values. Most of the values for the herd sensitivity are generated in a range between 0.9 and 1.0 for the Smith (1999) study.

Figure 9-4 depicts the conditional relationship between the median feedlot prevalence and the herd sensitivity for the Elder (2000) study. The graph implies that variation of the herd sensitivity does not have any effect on the median feedlot prevalence for this study. The median feedlot prevalence varies between approximately 65% and 100%, or a range of approximately 35 percentage points, depending on whether all other inputs are at their maximum or minimum values.

### 9.1.2 Application of Conditional Sensitivity Analysis to the Within Feedlot Prevalence Part

The within feedlot prevalence part in the production module is explained in Section 3.2.1 and inputs for this part are given in Table 3-9. The output of interest is the average within feedlot prevalence. There is a one-dimensional uncertainty simulation in this part as discussed in Section 3.3.1. Based upon the results of other sensitivity analysis methods, such as ANOVA and regression analysis, the study, apparent within feedlot prevalence, and test sensitivity were identified as the top three important inputs. Moreover, analysis with ANOVA clarified that the average within feedlot prevalence is higher during summer in comparison with winter. In

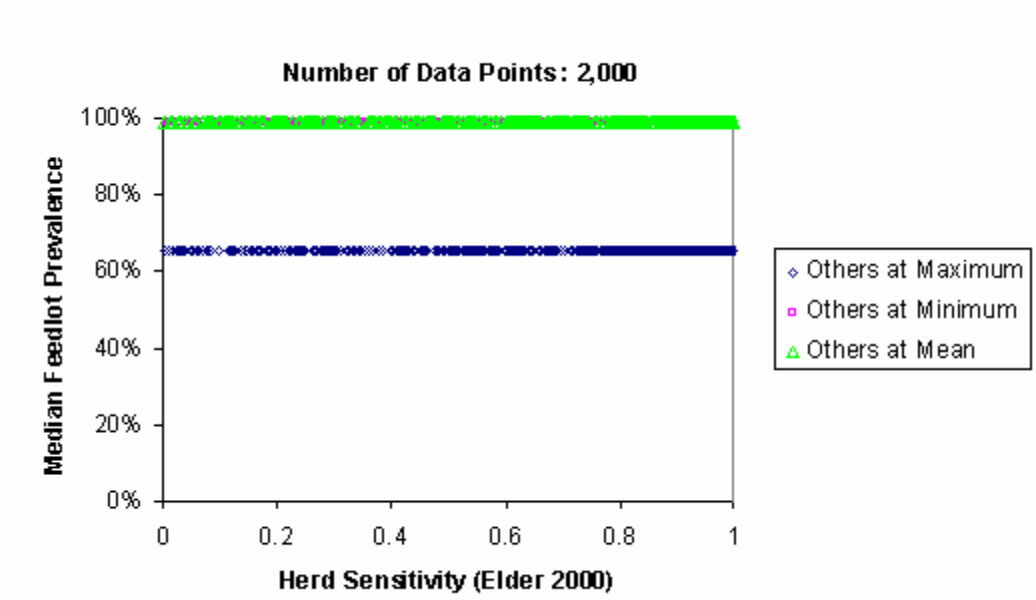



Figure 9-4. Conditional Sensitivity Analysis of the Herd Sensitivity, Elder 2000 Study.

Table 9-2. Nominal Values for the Apparent Within Feedlot and the Test Sensitivity in the Within Feedlot Prevalence Part

| Variable | Study | Minimum | Mean | Maximum | Unit |
|---|---|---|---|---|---|
| Apparent Within Feedlot Prevalence | Dargatz, Hancock 1997 | 0 | 2.8 | 100 | Percent |
| | Hancock 1999 | 0 | 2.5 | 100 | Percent |
| | Hancock 1998 | 0 | 3.7 | 100 | Percent |
| | Smith 1999 | 0 | 23 | 100 | Percent |
| | Elder 1999 | 0 | 36 | 100 | Percent |
| Test Sensitivity | Dargatz, Hancock 1997 | 0 | 58 | 100 | Percent |
| | Hancock 1999 | 0 | 58 | 100 | Percent |
| | Hancock 1998 | 0 | 58 | 100 | Percent |
| | Smith 1999 | 0 | 96 | 100 | Percent |
| | Elder 1999 | 0 | 96 | 100 | Percent |

addition, Table 3-2 indicates that Dargatz Hancock (1997) and Smith (1999) have the highest weight among the studies. Thus, for purposes of developing conditional sensitivity plots, the average within feedlot prevalence was plotted versus apparent within feedlot prevalence and test sensitivity for these two studies during the high prevalence season considering other inputs conditioned at minimum, mean, and maximum values. The nominal values for each input were

derived based on the input distribution. The nominal values for inputs to this part are given in Table 9-2.

Although there are two specific distributions for test sensitivity and apparent within feedlot prevalence given in Table 3-9, the parameters of these distributions are function of the study features such as number of samples in feedlots. Hence, nominal values for these inputs differ for each study.

For the conditional sensitivity analysis, for each study three simulations with 2,000 iterations were performed. In each simulation, the apparent within feedlot prevalence for that study was varied based on its distribution, while other inputs were conditioned at minimum, mean and maximum values. For the test sensitivity, apparent within feedlot prevalence was conditioned at its nominal values during the simulations, while the test sensitivity was allowed to vary based on its distribution. For the test sensitivity there are two testing methods: (1) "0.1g SMACct"; and (2) "10g IMS.

Figure 9-5 depicts the conditional relationship between the average within feedlot prevalence and the test sensitivity for the "0.1g SMACct" testing method. There is a nonlinear response to the variation of the test sensitivity when other inputs are conditioned at mean value. In this case, average within feedlot prevalence varies between approximately 8% and 20%, or a range of approximately 12 percentage points when the test sensitivity varies between 0.32 and 0.8. When other inputs are conditioned at maximum or minimum values, the average within feedlot prevalence remains constant at 100 and zero percent, respectively, with respect to the variation of the test sensitivity.

Figure 9-6 depicts the conditional relationship between the average within feedlot prevalence and the test sensitivity for "10g IMS" testing method. The conditional sensitivity graph in this figure implies that there is approximately a linear response to the variation of the test sensitivity when other inputs are conditioned at their mean values. In this case, average within feedlot prevalence varies between approximately 20% and 30%, or a range of approximately 10 percentage points when the test sensitivity varies between 0.7 and 1.0. When other inputs are conditioned at maximum or minimum values, the average within feedlot prevalence remains constant at 100 and zero percent, respectively, with respect to the variation of the test sensitivity.

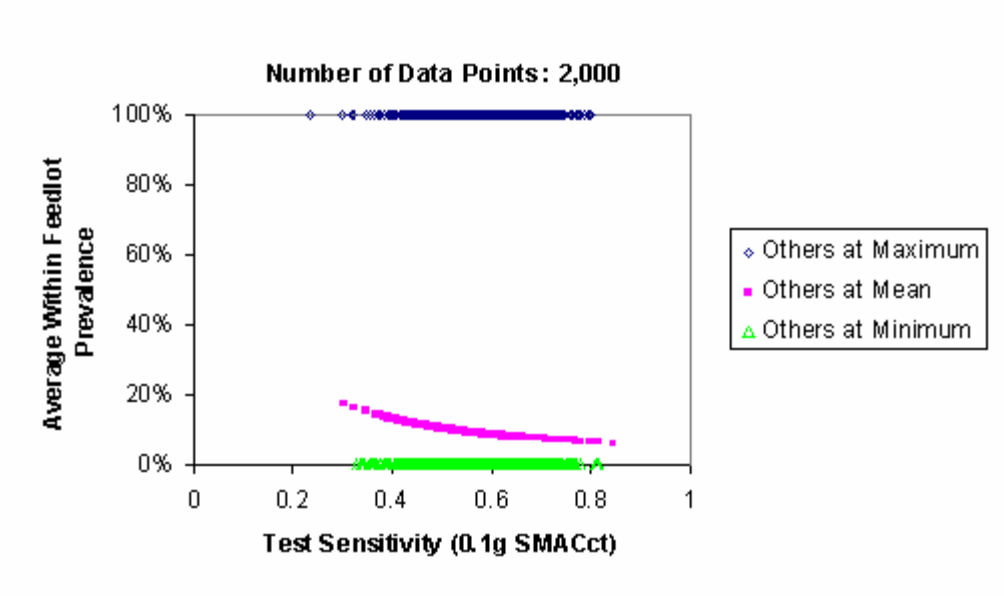



Figure 9-5. Conditional Sensitivity Analysis of the Test Sensitivity, 0.1g SMACct Testing Method.

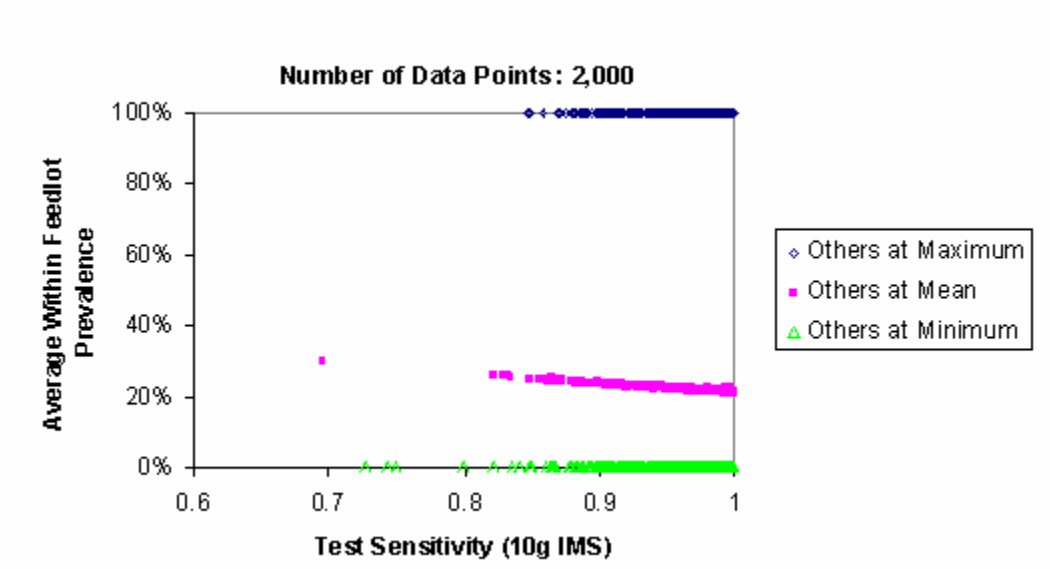



Figure 9-6. Conditional Sensitivity Analysis of the Test Sensitivity, 10g IMS Testing Method.

Figure 9-7 depicts the conditional relationship between the average within feedlot prevalence and apparent within feedlot prevalence for Dargatz Hancock (1997) study. The conditional sensitivity graph in this figure implies that there is approximately a linear response to the variation of the apparent within feedlot prevalence. When other inputs are conditioned at

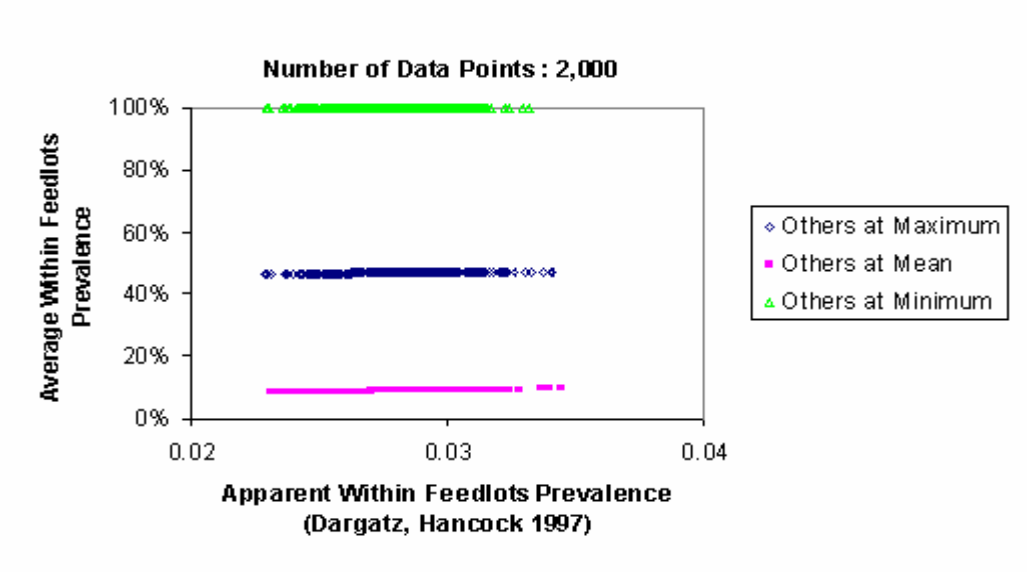



Figure 9-7. Conditional Sensitivity Analysis of the Apparent Within Feedlots Prevalence, Dargatz, Hancock 1997 Study.

minimum values, average within feedlot prevalence remains constant at 100 percent with respect to the variation of the apparent within feedlot prevalence.

Figure 9-8 depicts the conditional relationship between the average within feedlot prevalence and apparent within feedlot prevalence for Smith (1999) study. The conditional sensitivity graph in this figure implies that there is a linear response to the variation of the apparent within feedlot prevalence. When other inputs are conditioned at minimum values, average within feedlot prevalence remains constant at 100 percent with respect to the variation of the apparent within feedlot prevalence.

## 9.2 Application of Conditional Sensitivity Analysis to the Slaughter Module

Section 3.2.2 explains the slaughter module in the *E. coli* model. Inputs to the slaughter module are summarized in Table 3-10. The output of interest is the contamination in combo bins. The slaughter module includes both variability and uncertainty simulations. For simplicity, for conditional sensitivity analysis the variability only simulation was used. The case study scenario for the slaughter module is focused upon steers and heifers in the high prevalence season.

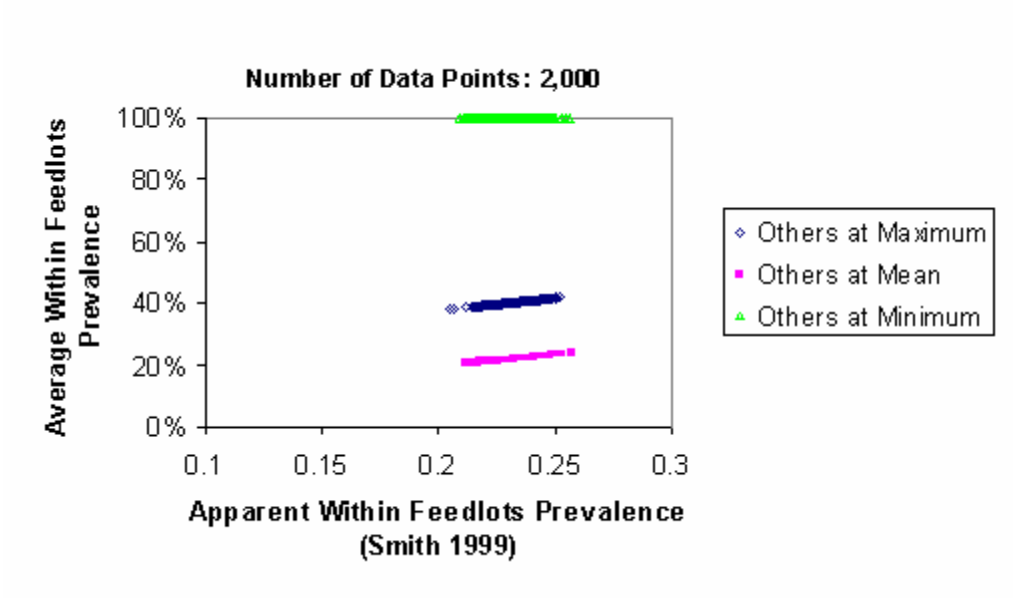



Figure 9-8. Conditional Sensitivity Analysis of the Apparent Within Feedlot Prevalence, Smith 1999 Study.

Table 9-3. Nominal Values for Inputs to the Slaughter Module

| Variable | Minimum | Mean | Maximum | Unit |
|---|---|---|---|---|
| Total Number of Combo Bin for Each Carcass | 2 | 4 | 6 | --- |
| Total Number of Infected Animals | 0 | 34 | 117 | --- |
| Total Number of Contaminated Animals | 0 | 54 | 117 | --- |
| Probability of Positive Cases at 2 Steps | 0 | 0.5 | 1 | --- |
| Number of Positive Cases at 2 Steps | 0 | 0 | 2 | --- |
| Number of Positive Cases at Evisceration | 0 | 0 | 2 | --- |
| Chilling Effect | -1 | 0 | 2.5 | Log |
| Number of Organisms | 0 | 9 | 3500 | --- |
| Trim Vacuum Washing Efficiency | 0 | 68 | 98 | Percent |
| Evisceration Organisms Added | 0 | 9 | 3500 | --- |
| Washing Effect | 0 | 90 | 99 | Percent |
| Contaminated $cm^2$ | 0 | 115 | 5600 | --- |

Based upon results with other sensitivity analysis methods, chilling effect, number of organisms, Trim/Vacuum/Wash efficiency, and washing effect were identified as the most important inputs and are the focus of analysis using conditional sensitivity. For purposes of developing conditional sensitivity plots, the combo bin contamination was plotted versus chilling

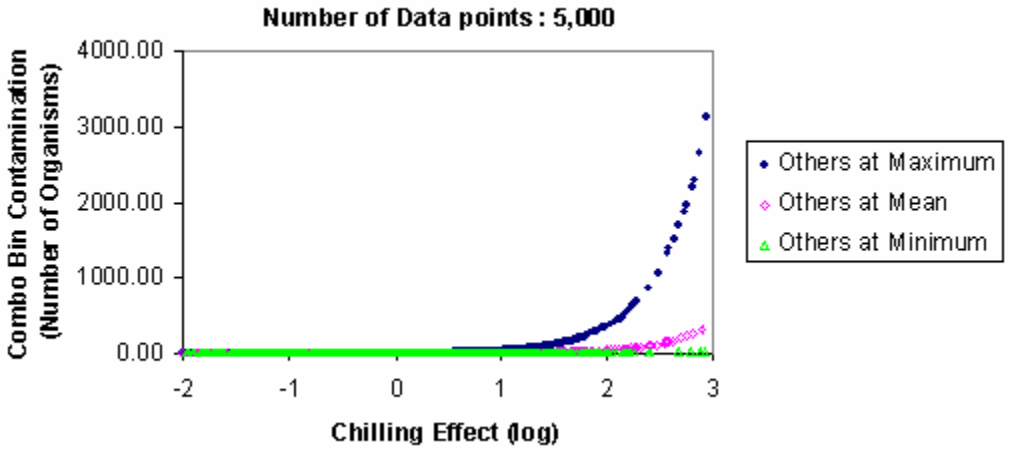



Figure 9-9. Conditional Sensitivity Analysis of the Chilling Effect.

effect, number of organisms, Trim/Vacuum/Wash efficiency, and washing effect considering other inputs conditioned at minimum, mean, and maximum values. The nominal values for each input were derived based on the input distribution. The nominal values for inputs to this part are given in Table 9-3.

In order to apply the conditional sensitivity analysis to each input, a simulation with 5,000 variability iterations was performed. In each case, a selected input was varied based on its distribution, while other inputs were conditioned at minimum, mean or maximum values.

Figure 9-9 presents conditional sensitivity plot for combo bin contamination versus the chilling effect. Based on the temperature during the chilling process the number of *E. coli* organisms on carcasses might increase or decrease (FSIS, 2001). This figure indicates that there is a nonlinear response to the variation of the chilling effect when other inputs are conditioned at their mean or maximum values. A large amount of combo bin contamination corresponds to cases for which the chilling effect is greater than an apparent threshold. For a chilling effect of less than 1 log there are approximately no *E. coli* organisms in combo bins when other inputs are conditioned at mean or maximum values. Moreover, when other inputs are conditioned at mean or maximum values, there is high amount of contamination (i.e., more than 2 logs or 100 organisms) when the chilling effect is larger than 2.5 logs. When other inputs are held at

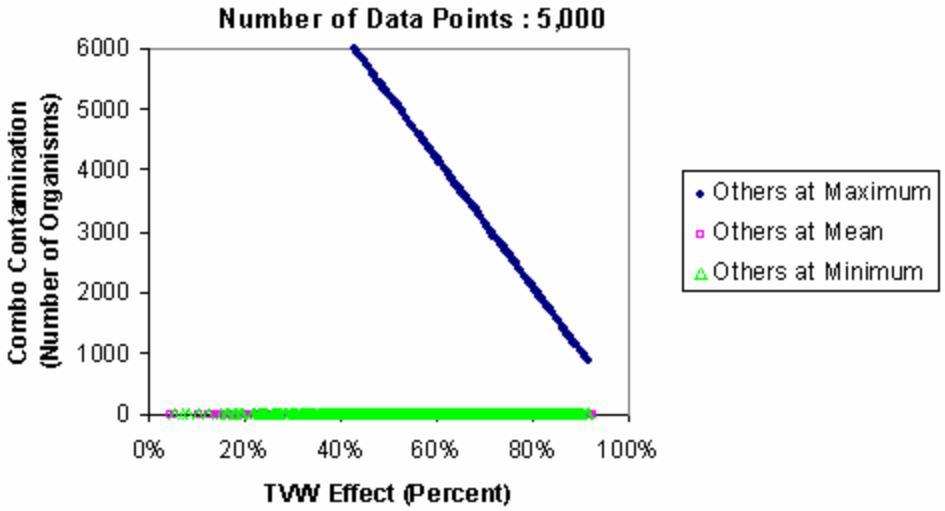



Figure 9-10. Conditional Sensitivity Analysis of the Trim/Vacuum/Wash Effect.

minimum values there will be no contamination in combo bins even if there is a high value of the chilling effect.

Figures 9-10 presents conditional sensitivity plot for combo bin contamination versus the Trim/Vacuum/Wash efficiency. There is a linear response to the variation of the Trim/Vacuum/Wash efficiency. When other inputs are conditioned at their maximum values, the contamination levels are substantial even at the highest possible Trim/Vacuum/Wash efficiency. This indicates that performing the decontamination step with the highest efficiency may not guarantee a low value of combo bin contamination. However, when other inputs are at their mean or minimum values, the contamination level is relatively insensitive to the Trim/Vacuum/Wash efficiency. These results imply that Trim/Vacuum/Wash efficiency is important only if other inputs are at sufficiently high values. Thus, there is an interaction between Trim/Vacuum/Wash efficiency and other inputs.

Figures 9-11 presents a conditional sensitivity plot for combo bin contamination versus the washing effect. The washing effect presents the same pattern as Trim/Vacuum/Wash efficiency in Figure 9-10. There is a linear response to the variation of the washing effect. When other inputs are conditioned at their maximum values, the contamination levels are substantial even at the highest possible washing efficiency. This indicates that even the highest efficiency during the decontamination step using washing may not guarantee a low value of combo bin

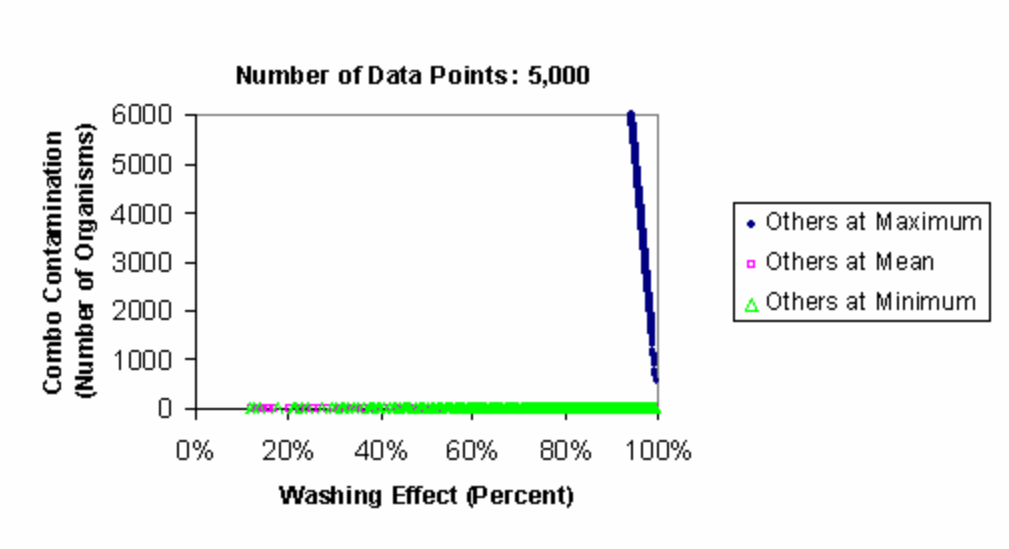



Figure 9-11. Conditional Sensitivity Analysis of the Washing Effect.

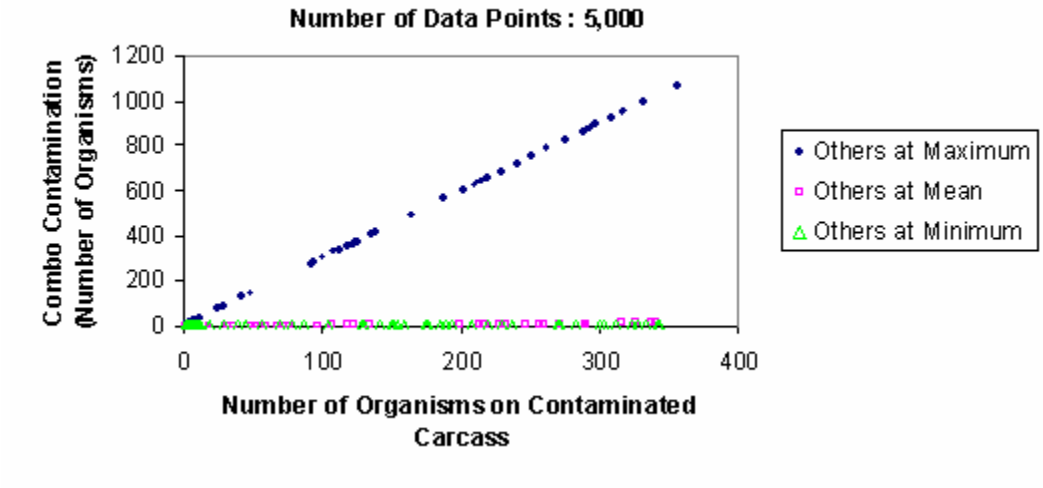



Figure 9-12. Conditional Sensitivity Analysis of the Number of Organisms.

contamination. However, when other inputs are at their mean or minimum values, the contamination level is relatively insensitive to the washing efficiency. These results imply that washing efficiency is important only if other inputs are at sufficiently high values. Thus, there is an interaction between this input and other inputs.

Figures 9-12 presents a conditional sensitivity plot for combo bin contamination versus the number of organisms on contaminated carcass. When other inputs are conditioned at maximum values, there is a linear relationship between the combo bin contamination and the number of organisms on contaminated carcasses. When other inputs are conditioned at their

Table 9-4. Nominal Values for Inputs to the Growth Estimation Part

| Variable | Minimum | Mean | Maximum | Unit |
|---|---|---|---|---|
| Storage Time at Retail | 0 | 24 | 340 | Hour |
| Storage Temperature at Retail | 46 | 47.6 | 73 | $^{o}$F |
| Storage Time at Transportation | 0 | 1 | 6.5 | Hour |
| Storage Temperature at Transportation | 46 | 48.8 | 73 | $^{o}$F |
| Storage Time at Home | 0 | 24 | 340 | Hour |
| Storage Temperature at Home | 46 | 48.3 | 73 | $^{o}$F |
| Maximum Density | 5 | 7.5 | 10 | log |

mean or minimum values there is no combo contamination regardless of the number of organisms. These results imply that there is an interaction between the number of organisms on contaminated carcasses and other inputs.

## 9.3 Application of Conditional Sensitivity Analysis to the Preparation Module

In the preparation module conditional sensitivity analysis was applied to two parts, including the growth estimation and the serving contamination parts. Because the relationship between the output and inputs in the cooking effect part is pre-defined in the form of linear models for each precooking treatment, application of the conditional sensitivity analysis to the cooking effect part is not informative. Thus, the conditional sensitivity plots are presented for the growth estimation and serving contamination parts in Sections 9.3.1 and 9.3.2, respectively.

### 9.3.1 Application of Conditional Sensitivity Analysis to the Growth Estimation Part

The growth estimation part in the preparation module is explained in Section 3.2.3 and inputs for this part are given in Table 3-11. The output of interest is the mean growth of the *E. coli* organisms in ground beef servings. The growth estimation part includes both variability and uncertainty simulations. For simplicity, for conditional sensitivity analysis in this part variability only simulation was considered by holding all uncertain inputs at their point estimates.

Based upon results with other sensitivity analysis methods, the storage time and the storage temperature at stages 1 and 3 were identified as the most important inputs and are the focus of analysis using conditional sensitivity. For purpose of developing conditional sensitivity plots, the mean growth in the ground beef servings was plotted versus the storage temperature and the storage time at stages 1 and 3 considering other inputs conditioned at minimum, mean, and maximum values. The nominal values for each input were derived based on their input distributions. The nominal values for inputs to this part are given in Table 9-4.

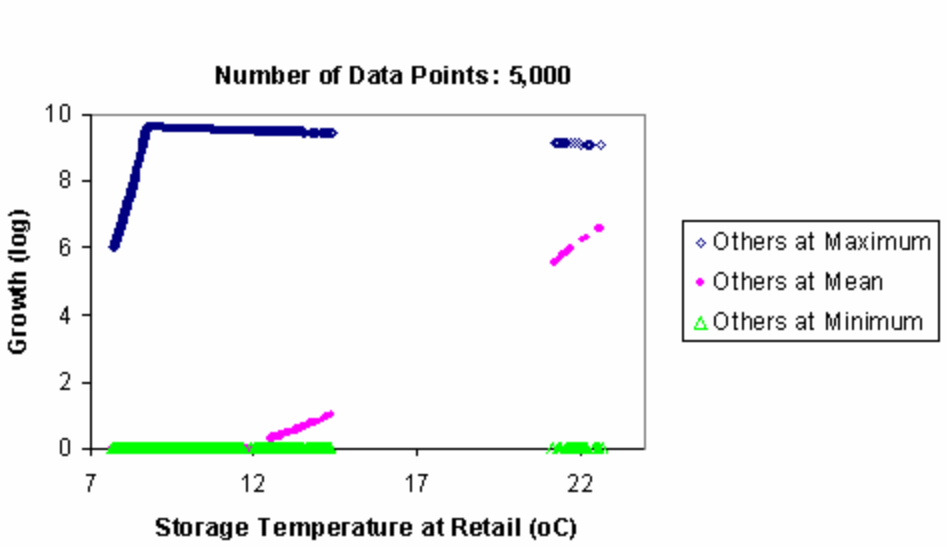



Figure 9-13. Conditional Sensitivity Analysis of the Storage Temperature at Retail.

In order to apply the conditional sensitivity analysis to each input, a simulation with 5,000 variability iterations was performed. In each case, the selected input was varied based on its distribution, while other inputs were conditioned at minimum, mean and maximum.

Figure 9-13 presents a conditional sensitivity plot for the mean growth versus the storage temperature at stage 1. When other inputs are held at their mean values, there is no growth unless the storage temperature at stage 1 is greater than approximately 12$^{o}$C. In this case, when the storage temperature increases this value, there is a nonlinear response to the increase of the storage temperature. The approximate value of 12 $^{o}$C can be considered as a threshold in the model response to the variation of the storage temperature at stage 1, when other inputs are conditioned at their mean values. If other inputs are held at their maximum values there is a large amount of growth for ground beef servings even at low temperature. With increase in the storage temperature the growth in the ground beef serving increases nonlinearly until the saturation point for the growth of *E. coli* organisms is reached at temperature of approximately 8.7$^{o}$C. After this temperature there is a nonlinear decrease in the maximum possible growth of the *E. coli* organisms. The decrease in the estimated growth after reaching the saturation point is because of the decrease in the maximum population density factor. The maximum population density is function of the storage temperature and it decreases with increase in the storage temperature (FSIS, 2001). In addition, there is a gap in Figure 9-13. The storage temperature at stage 1

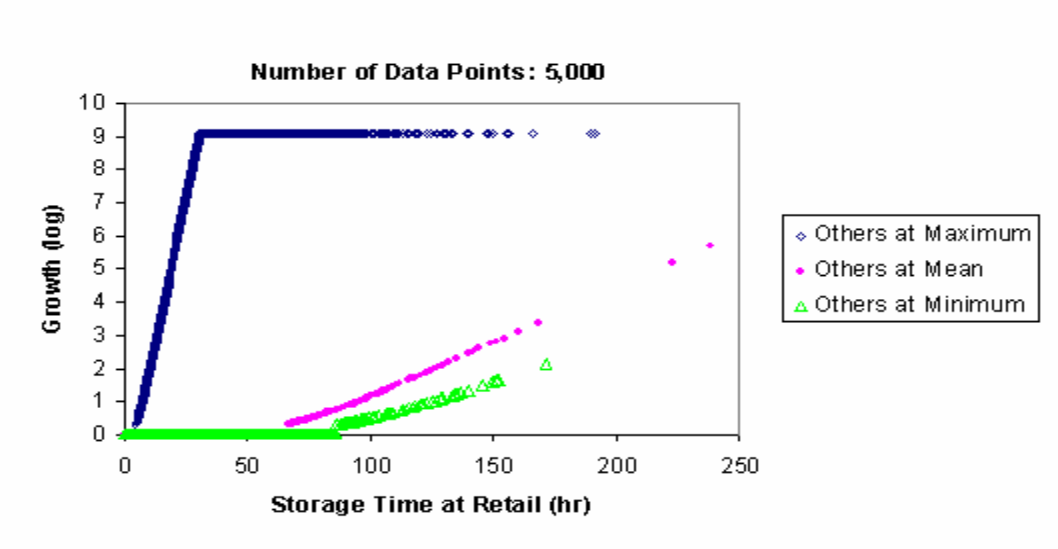



Figure 9-14. Conditional Sensitivity Analysis of the Storage Time at Retail.

between 14°C and 21°C were not generated in the random simulation. This is likely attributable to an error in the original model.

Figure 9-14 presents a conditional sensitivity plot for the mean growth versus the storage time at stage 1. 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 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 for the growth in the ground beef servings due to the storage time at stage 1 when other inputs are conditioned at their minimum values is higher with an approximate value of 86 hrs indicating that there is an interaction between the storage time and other inputs. If other inputs are held at their maximum values the threshold for the growth of the *E. coli* organisms is approximately 4.5 hrs. When other inputs are conditioned at their maximum values the saturation point for the growth of *E. coli* organisms is reached after approximately 31 hrs. In contrast, when other inputs are held at their minimum or mean values the saturation point is not achieved even with storing the ground beef servings for a long time.

Figure 9-15 presents a conditional sensitivity plot for the mean growth versus the storage temperature at stage 3. When other inputs are held at their mean values, there is no growth unless the storage temperature at stage 3 of approximately 12.7°C. In this case, there is a nonlinear

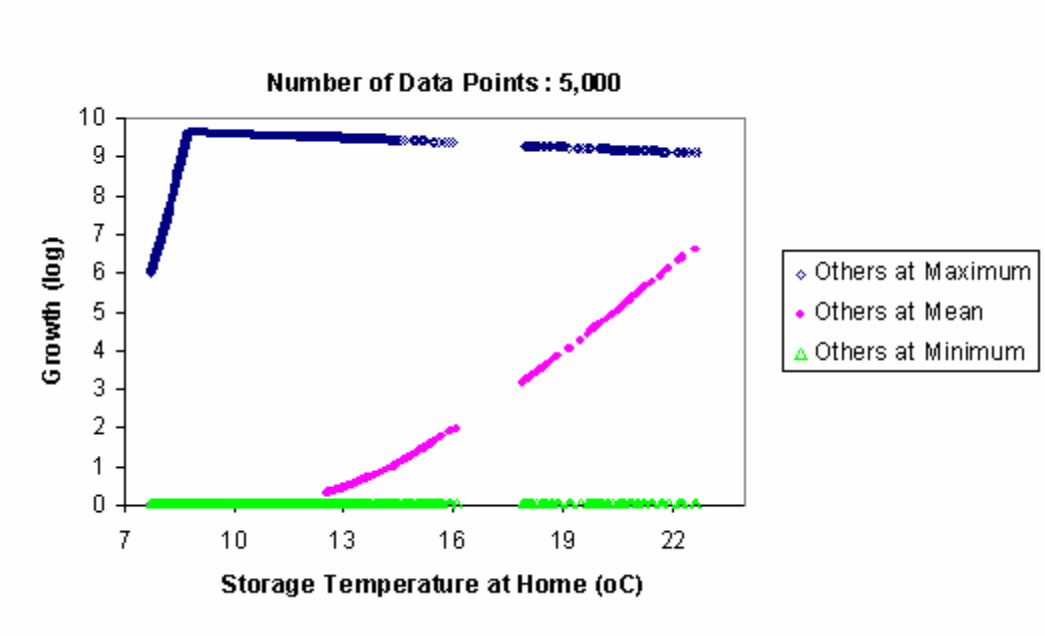



Figure 9-15. Conditional Sensitivity Analysis of the Storage Temperature at Home.

response to the increase of the storage temperature above this value. The approximate value of 12.7 $^{o}$C can be considered as a threshold in the model response to the variation of the storage temperature at this stage, when other inputs are conditioned at their mean values. If other inputs are held at their maximum values there is a large growth rate for ground beef servings even at low storage temperatures. With an increase in the storage temperature the growth in the ground beef serving increases nonlinearly until the saturation point for the growth of *E. coli* organisms is reached at temperature of approximately 8.7$^{o}$C. Above this temperature there is a nonlinear decrease in the maximum possible growth of the *E. coli* organisms. In addition, there is a gap in Figure 9-13. The storage temperature at stage 3 between 16$^{o}$C and 18$^{o}$C were not generated in the random simulation. This is likely attributable to an error in the original model.

Figure 9-16 presents a conditional sensitivity plot for the mean growth versus the storage time at stage 3. When other inputs are held at their mean values, there is no growth unless the storage time is greater than approximately 64 hrs. In this case, when the ground beef servings are stored more than 64 hrs, there is a nonlinear response to the increase of the storage time. The approximate value of 64 hrs can be considered as a threshold in the model response to the variation of the storage time at home, when other inputs are conditioned at their mean values. The threshold for the growth in the ground beef servings due to the storage time at stage 3 when other inputs are conditioned at their minimum values is higher with an approximate value of 88

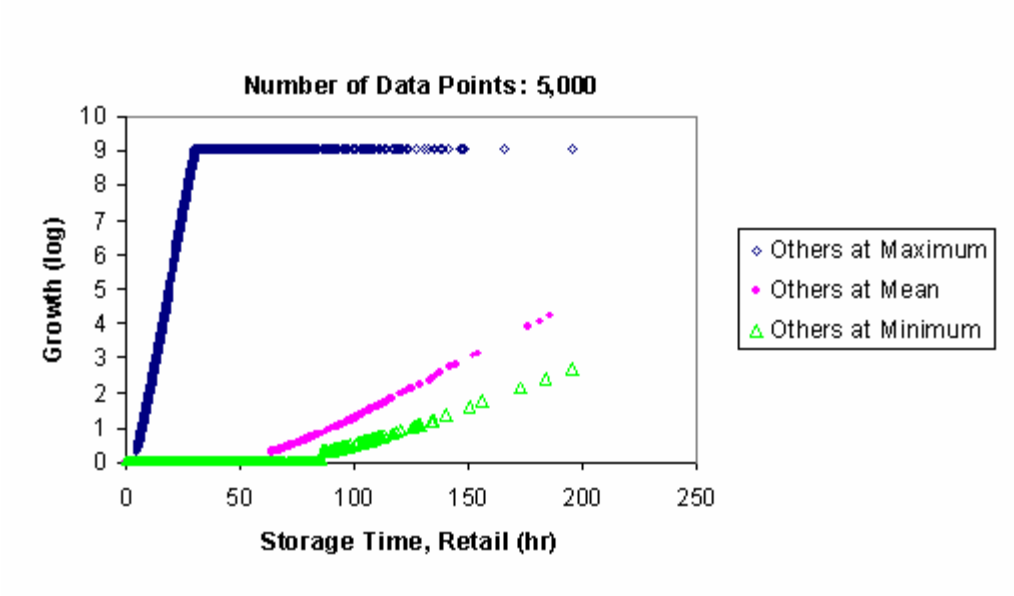



Figure 9-16. Conditional Sensitivity Analysis of the Storage Time at Home.

hrs. The difference in thresholds 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 5.1 hrs. When other inputs are conditioned at their maximum values the saturation point for the growth of *E. coli* organisms is reached after approximately 29 hrs. In contrast, when other inputs are held at their minimum or mean values the saturation point is not achieved even with storing the ground beef servings for a long time.

### 9.3.2 Application of the Conditional Sensitivity Analysis to the Serving Contamination Part

The serving contamination part in the preparation module is explained in Section 3.2.3 and inputs to this part are given in Table 3-12. The output of interest is the mean serving contamination. There is a one-dimensional variability simulation in this part as discussed in Section 3.3.3. Conditional sensitivity graphs are prepared for the grinder contamination and the serving size in this part.

For purpose of developing conditional sensitivity plots, the mean serving contamination was plotted versus the grinder contamination or the serving size considering other inputs conditioned at their minimum, mean, or maximum values. Most of the inputs in the serving contamination part, such as the ground beef consumption type, the eating location, and the consumer age are qualitative. Hence, for these inputs it is not possible to define nominal values.

Table 9-5. Nominal Values for the Grinder Contamination and the Serving Size in the Serving Contamination Part

| Variable | Min | Mean | Max | Unit |
|---|---|---|---|---|
| Grinder Contamination from Combo Bins | -7 | -4 | -1 | Log |
| Grinder Contamination from Trim Boxes | -7 | -6 | -2 | Log |
| Serving Size for Hamburger Patties at Home | 5.1 | 105 | 448 | g |
| Serving Size for Hamburger Patties Away | 15 | 90 | 500 | g |

Note: The nominal values for the serving size are derived from the table presented in the *E. coli* model and in "CUNSUMPTION" worksheet. The source of the data is CFSII 1994-1996, 1998.

Therefore, for these inputs specific levels were selected based on the results from other sensitivity analysis methods. For example, Section 5.4.3 presented results of different contrasts in the serving contamination part. Based on the contrasts results summarized in Table 5-26, servings consumed by people between 25 and 64 years old have higher risk of contamination. Hence, this age group is considered for the consumer age level. Moreover, Figure 3-16 indicates that the hamburger patties have the highest amount of consumption in the United States. Thus, hamburger is selected for the ground beef consumption type level. FSIS (2001) indicates that the high prevalence season has higher risk of contamination in ground beef servings. Therefore, summer is selected for the conditional sensitivity analysis. For quantitative inputs, nominal values are given in Table 9-5.

In Table 9-5 the nominal values for the grinder contamination and the serving size are summarized. The nominal values for the serving size are specific for hamburger patties. In order to apply the conditional sensitivity analysis to each input, a simulation with 5,000 variability iterations was performed. In each case, the selected input was varied based on its distribution, while other inputs were conditioned at minimum, mean or maximum.

Figures 9-16 and 9-17 present conditional sensitivity plots for serving contamination versus the grinder contamination at home and away from home, respectively. These figures indicate that there are nonlinear responses to the variation of the grinder contamination when other inputs are conditioned at their nominal values. When serving size is conditioned at its mean value, contamination in hamburger patties is negligible when grinder loads have contamination less than approximately -2.5 logs. Therefore, this value can be considered as a threshold in the model response to the variation of the grinder. A comparison of Figures 9-16 and 9-17 provide insights regarding whether contamination can be greater at home or away. For example, when all other inputs are conditioned at their mean values, the serving contamination at home is larger

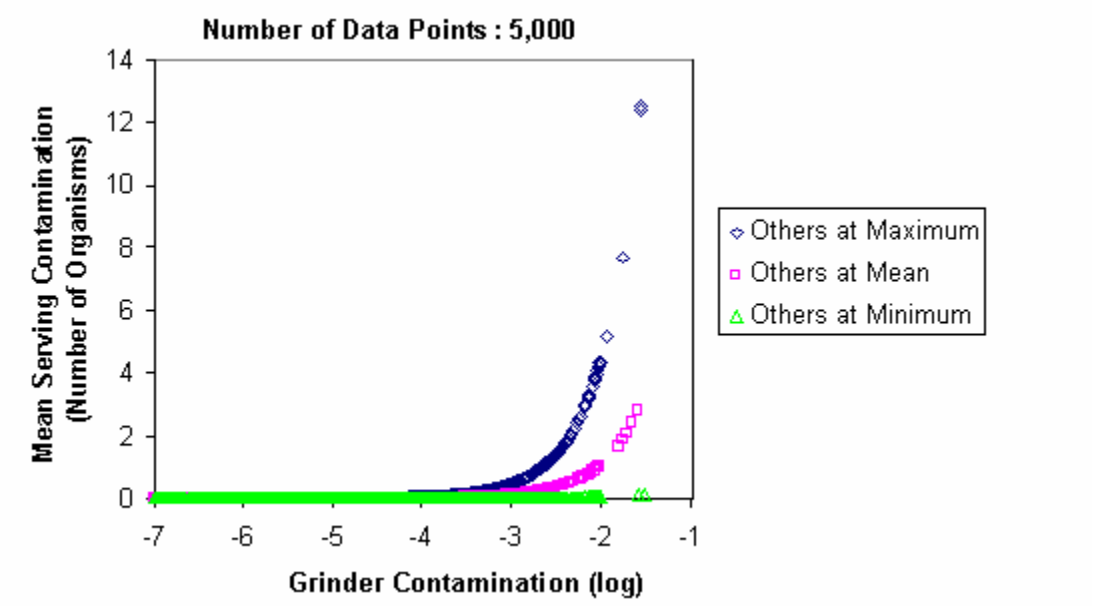



Figure 9-17. Conditional Sensitivity Analysis of the Grinder Contamination Effect at Home.

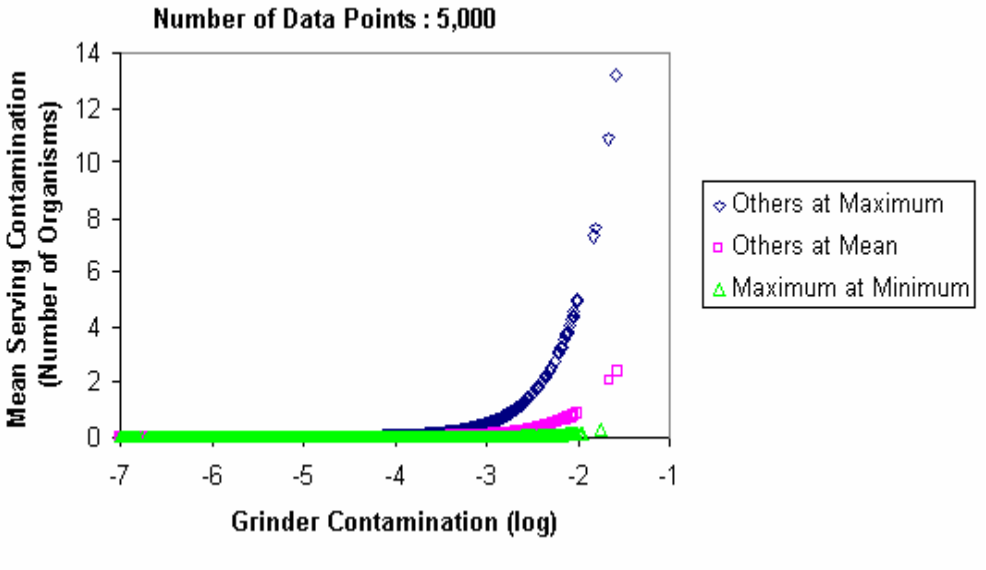



Figure 9-18. Conditional Sensitivity Analysis of the Grinder Contamination Effect Away from Home.

than that away from home for a given grinder contamination. This pattern was also identified using contrasts in ANOVA in Section 5.4.3.

Figures 9-18 and 9-19 present conditional sensitivity plots for serving contamination versus the serving size for hamburger patties consumed at home and away from home, respectively. These figures indicate that there are linear responses to the variation of the serving

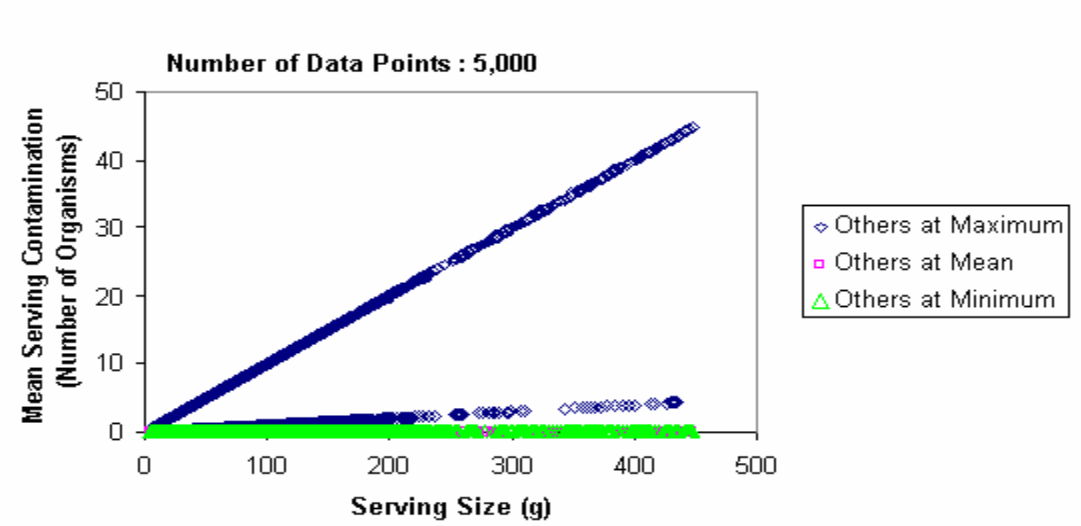



Figure 9-19. Conditional Sensitivity Analysis of the Serving Size Effect at Home.

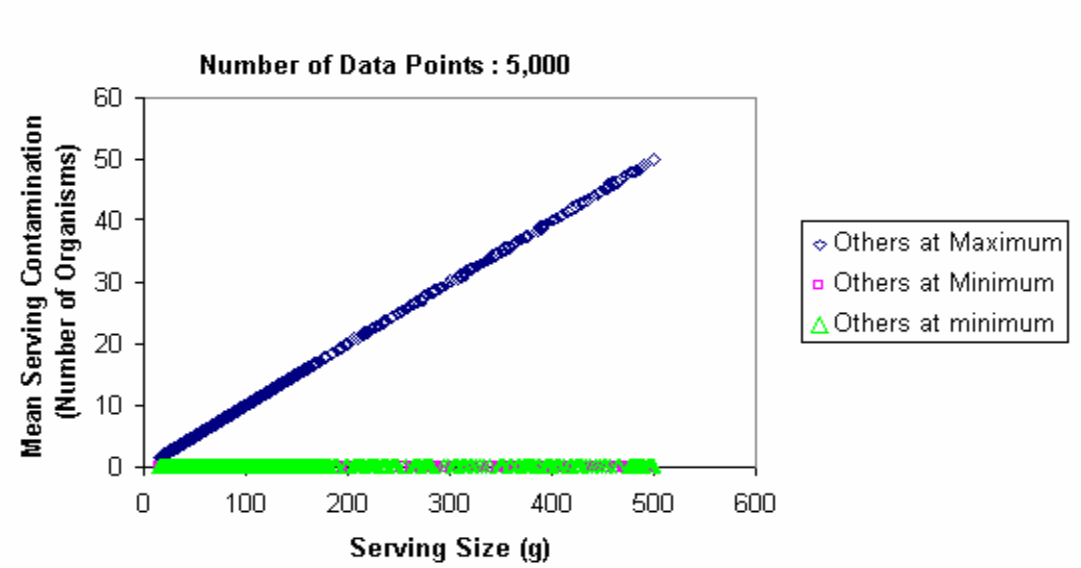



Figure 9-20. Conditional Sensitivity Analysis of the Serving Size Effect Away from Home.

size when other inputs are conditioned at their maximum values. Hence, with increase in the serving size higher contamination in hamburger patties is expected. In Figure 9-18 two lines are depicted for the case when other inputs are conditioned at their maximum values. This occurs because for the servings consumed at home, the grinder load has two different sources of meat trims. These include meat trims coming from combo bins and meat trims coming from trim boxes. Grinder loads filled with meat trims from trim boxes have lower contamination (FSIS, 2001). The line with the lower slope is for meat trims coming from trim boxes. However, when

other inputs are at their mean or minimum values, the serving contamination level is relatively insensitive to the serving size. These results imply that serving size is important only if other inputs are at sufficiently high values. Thus, there is an interaction between the serving size and other inputs.

## 9.4 Evaluation of Conditional Sensitivity Analysis Based on Applications to the *E. coli* Model

In Sections 9-1 to 9-3 conditional sensitivity analysis was applied to different modules and parts of the *E. coli* model. Conditional sensitivity analysis was implemented in order to clarify special relationship between the output and inputs such as non-linearity in the model response, thresholds, and interactions. Conditional sensitivity plots cannot be used to explicitly rank the inputs. However, the possibility of clarifying special relationships such as those mentioned above is as an advantage of this method.

Non-linearity in the model response to specific inputs can be identified using conditional sensitivity analysis. For example, in Section 9.3.2 there is a non-linear response of the model to the variation of the grinder contamination. This trend cannot easily be identified by some sensitivity analysis method such as linear regression analysis, which assumes a specific functional relationship between the output and each input.

Thresholds in the model response to a specific input can be identified using conditional sensitivity analysis. For example, in Section 9.3.2 a threshold in the model response to the grinder contamination was identified. The capability to identify thresholds with conditional sensitivity analysis is useful in comparison to other methods, such as CART and ANOVA. This point was also discussed for scatter plots in Section 8.4.

Conditional sensitivity analysis can be used in order to clarify interaction effects between inputs. For example, in Section 9.3.1 a conditional sensitivity plot identified that there is an interaction between the storage time and other inputs.

Conditional sensitivity analysis method can also be used in order to verify the model structure. For example, Figures 9-13 and 9-15 depict the conditional sensitivity plots for the mean growth in ground beef servings versus the storage temperature at retail and home. In these figures there are gaps in graphs indicating that the growth estimation model cannot generate the storage temperatures between $14^{o}C$ and $21^{o}C$, and between $16^{o}C$ and $18^{o}C$ for retail and home, respectively. This is likely attributable to an error in the original model.

The examples presented in this chapter were conditioned on minimum, mean, and maximum values of all other inputs. Of course, the likelihood that all other inputs would simultaneously take on their minimum values is rare. Similarly, it is rare that all of the inputs would simultaneously take on their mean or maximum values. Thus, it is not possible to directly infer the relative importance of different types of model responses conditioned on the arbitrary assumptions made regarding the other model inputs. However, the key qualitative insight that this method affords is regarding whether nonlinearities, thresholds, and interactions exist and their characteristics if they do. This information is valuable in choosing other sensitivity analysis methods and in targeting additional analysis to further clarify and explore the significance of such relationships.

# 10 SENSITIVITY IN EXPOSURE ASSESSMENT IN GROUND BEEF SERVINGS

The objective of this chapter is to identify the priority order of key modules and parts of the *E. coli* model explained in Chapter 3 in order to attain a general insight about the relative importance of different parts of the model.

As explained in Section 3.2.4, one of the most important goals of sensitivity analysis, as described at the NCSU/USDA Workshop on Sensitivity Analysis, is to perform global sensitivity analysis on output variables of direct relevance to a decision. Global sensitivity analysis in food safety risk assessment models facilitates exploring effective approaches for mitigation of morbidity and mortality risk of the food-born pathogen. However, as explained in Section 3.2.4.1, because the *E. coli* model is divided into modules global sensitivity analysis cannot be performed. Figure 10-1 depicts the connection between the final output of the model in the exposure part and each module and part of the model.

Figure 10-1 depicts that outputs of one module serve as inputs to the next. In combination with the fact that many of the intermediate values of inputs are binned, the implication of both modularity and binning of variables is that there is a lack of one-for-one correspondence between the values of a desirable risk assessment model output, such as exposure to *E. coli* organisms in ground beef servings, and the values of inputs to the various modules that influence the output. This modeling structure forced the application of local sensitivity analyses in individual modules and parts of the model, as presented in Chapters 4 to 9. Although it is not practical to have a global sensitivity analysis, a question was raised about exploring a way that different modules and parts of the *E. coli* model could be prioritized based on their importance on the exposure to *E. coli* organisms. In order to answer to this question some background information for the final model output in the exposure part is presented in the following.

The amount of *E. coli O157:H7* to which a consumer might be exposed in a single serving of ground beef is a function of the original number of *E. coli O157:H7* organisms and the subsequent effects of storage, handling, and cooking on the growth or decline in the number of *E. coli* organisms in ground beef (FSIS, 2001). The original number of *E. coli* organisms is estimated based on information from production, slaughter, and preparation modules, while the effect of growth and decline in the number of organisms are estimated in the preparation module. The final dose distribution to which the population is exposed ($DOSE_{pop}$) is expressed as the

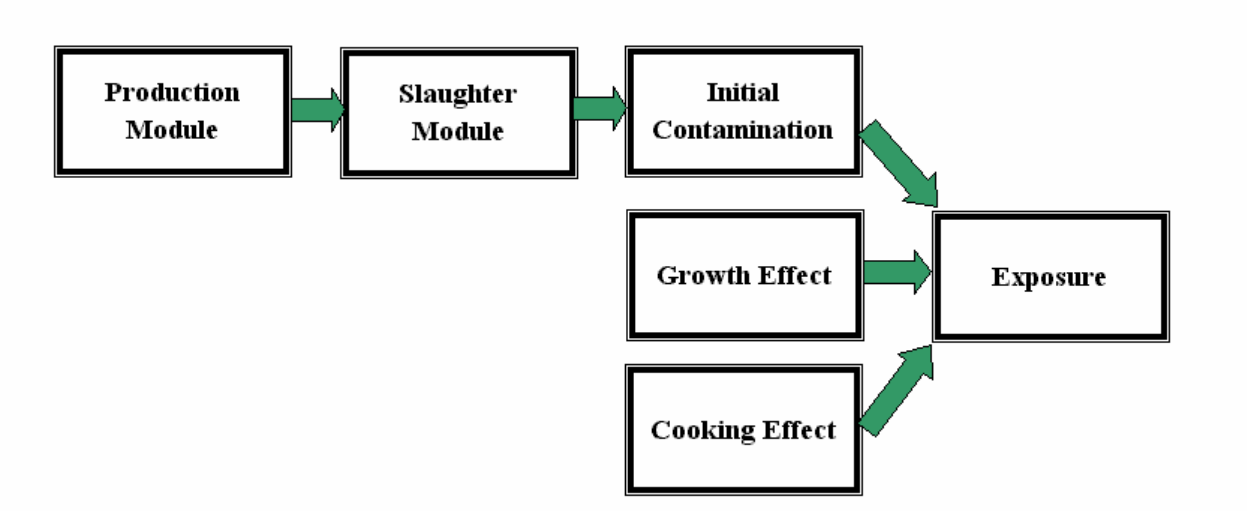



Figure 10-1. Schematic Diagram of the Connection Between the Exposure and Different Modules and Parts of the *E. coli* Model.

initial serving distribution plus the growth distribution minus the distribution describing the effect of cooking:

$$DOSE_{pop} = BACT_{pop} + G_{pop} - LR_{pop} \quad (10\text{-}1)$$

In above equation, $BACT_{pop}$ is the distribution for the initial number of *E. coli* organisms in servings, while $G_{pop}$ and $LR_{pop}$ are the effect of growth and cooking on the ground beef servings, respectively. In Section 10.1 a case study for evaluation the relative ranking of these three distributions is presented.

## 10.1 Case Study for Ranking the Factors Affecting the Final Exposure to *E. coli O157:H7* in Ground Beef Servings

A case study was prepared to prioritize the effect of different modules of the *E. coli* food model on the final exposure in ground beef serving. Equation 10-1 implies that for estimation of the final exposure distribution, three distributions for the initial concentration, the growth effect, and the cooking effect should be available. For estimation of these distributions, the *E. coli* model was set to run a random simulation of 5,000 variability and 50 uncertainty iterations. The initial concentration of the *E. coli* organisms in ground beef servings is available based on the output from the preparation module in the serving contamination part. The initial number of *E. coli* organisms is a function of the outputs estimated in the production part and the slaughter module such as the infection prevalence and combo bins contamination. Related distributions for

Table 10-1. Different Cases in Sensitivity Analysis of the Exposure Estimation Part

| | Initial Concentration | Growth Effect | Cooking Effect |
|---|---|---|---|
| **Case Zero** | Vary | Vary | Vary |
| **Case One** | Constant | Constant | Vary |
| **Case Two** | Constant | Vary | Constant |
| **Case Three** | Vary | Constant | Constant |

the growth and the cooking effect are also available from the intermediate outputs of the preparation modules in growth estimation and cooking effect parts.

In Figures 10-2 to 10-4 distributions for the *E. coli* initial concentration, the growth effect and the cooking effect are depicted, respectively. The case study focused on the high prevalence season for the analysis because of the higher risk of exposure to high levels of contamination in ground beef servings during summer. The initial contamination distribution in Figure 10-2 is the average of the 50 distributions estimated for each uncertainty iteration. Figures 10-3 and 10-4 depict the average growth and cooking effect for 50 uncertainty iterations, respectively.

In order to estimate the exposure based on Equation 10-1, a Monte Carlo simulation was used. Detailed information about this simulation is presented in Section 10-2. For evaluation of the priority rank of the distributions with respect to variability in exposure, four cases were examined. These cases are summarized in Table 10-1. Case *Zero* represents the scenario in which all three variables (i.e., initial concentration, growth effect, and cooking effect) vary based on their distributions. Case *One* is the situation in which the cooking effect varies based on its distribution, while the other two variables are conditioned at their mean values. Case *Two* is based upon variability in growth effect while the other two inputs are held at their mean values. Case *Three* is similar except that the initial concentration varies while the other two inputs are constant at their mean values.

## 10.2 Analysis of Significant Parts in the Exposure Assessment

Section 10-1 explained that for analysis of the significant parts in the exposure assessment, Monte Carlo simulations were used for the exposure distribution estimation based on Equation 10-1 considering four cases in Table 10-1. Figures 10-2 and 10-3 implied that the probability of positive logs of *E. coli* organisms was small. Therefore, in order to have randomly generated values in the right tail of these distributions large number of iterations was used. Thus, 600,000 Monte Carlo iterations were performed using @Risk. In Figure 10-5 the results of the

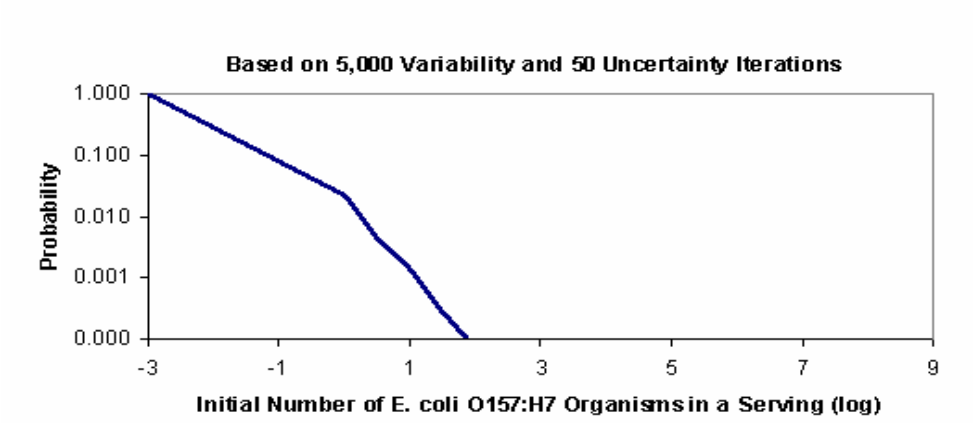



Figure 10-2. Probability Distribution Function for the Initial Number of *E. coli* Organisms in Ground Beef Servings in High Prevalence Season.

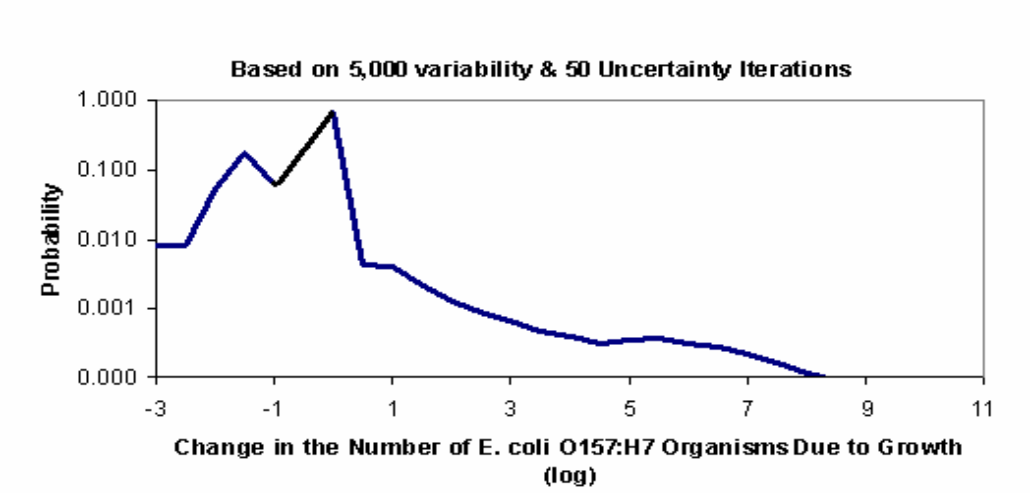



Figure 10-3. Probability Distribution Function for the Log Increase or Log Decrease Due to Storage of Ground Beef Servings.

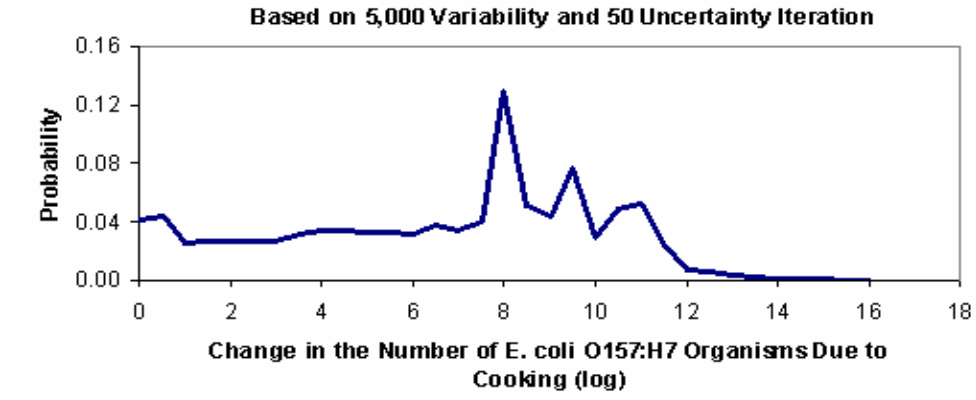



Figure 10-4. Probability Distribution Function for the Log Reduction in the Number of *E. coli* Organisms in Ground Beef Servings Due to Cooking.

analysis in the form of different variability distributions for the exposure to *E. coli* organisms in ground beef servings in summer are provided for cases *One, Two,* and *Three*.

Figure 10-5 indicates that Case *One*, representing the variation of only the cooking effect, leads to the widest range of variation in exposure. Cases *Two* and *Three* results in the second and third widest ranges. Therefore, the variability in exposure is more sensitive to the cooking effect than to growth effect or the initial concentration.

Risk managers might be interested in knowing which of the four cases have the largest maximum contamination or the highest probability of exceeding contamination levels that would be considered to be high. In order to address this interest, the maximum level of contamination generated in each of the four cases studies is summarized in Table 10-2. For Cases Zero, Two, and Three, the probabilities that the contamination levels reach the respective maximum values range from approximately $2\times10^{-6}$ to $8\times10^{-6}$. For Case One, four percent of the outcomes are at the maximum contamination value. Based upon the estimate of $1.823\times10^{10}$ annual ground beef servings in the U.S. (FSIS, 2001), the estimated number of servings that contain the respective maximum contamination levels are shown for the four cases. The value of the maximum contamination level differs in each of the four cases. The highest contamination level is associated with Case Zero, in which all three of the input distributions vary. Of the three cases in which only one distribution varies at a time, it is clear that Case Two has the largest value of the maximum contamination compared to the other two. Thus, it appears that the growth effect portion of the model contributes the most to the possibility of high estimated maximum concentration values.

Based on the presented analyses, the cooking effect part caused a maximum range of variation in the exposure to *E. coli* organisms and growth effect eventuated in the highest number of consumers exposed to positive dose of *E. coli* organisms in ground beef servings.

The cooking effect and the growth estimation parts were individually analyzed with several sensitivity analysis methods in Chapters 4 to 9 and important inputs to these parts were identified. For the cooking effect part, cooking temperature was selected as the most important input, while for the growth estimation time the storage time and the storage temperature at retail and home were identified as high ranked inputs. Hence, management strategies should be

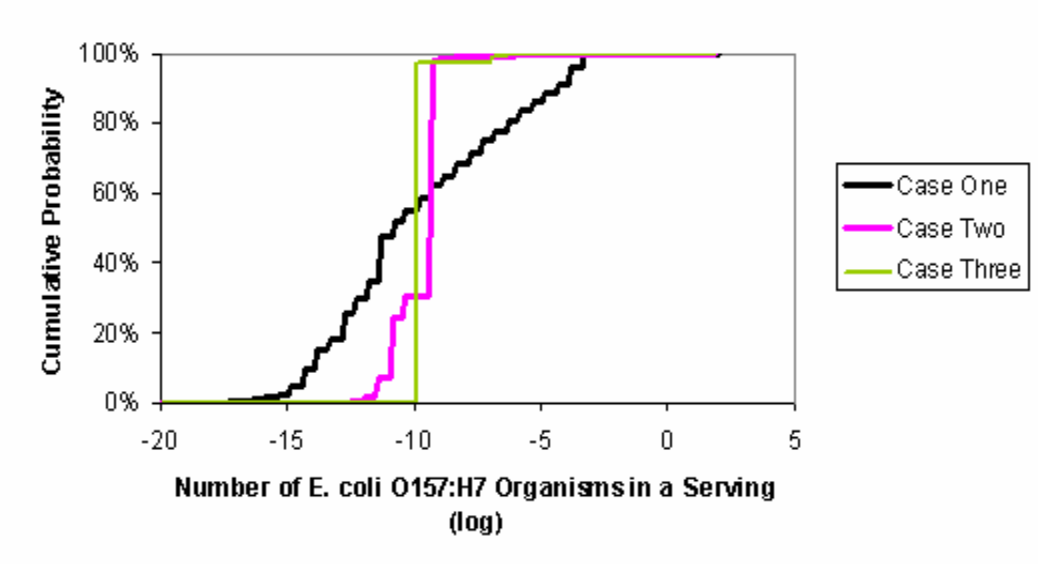



Figure 10-5. Exposure Assessment for the Ground Beef Servings in High Prevalence Season.

Table 10-2. Maximum Ground Beef Servings Contamination

| | Maximum Contamination (log) | Probability of Maximum Contamination | Number of Servings with Maximum Contamination |
|---|---|---|---|
| **Case Zero** | 7.9 | $1.67*10^{-6}$ | 30,440 |
| **Case One** | -3.5 | 0.04 | 729,128,260 |
| **Case Two** | 0.03 | $5*10^{-6}$ | 91,140 |
| **Case Three** | -4.5 | $8.3*10^{-6}$ | 151,290 |

focused on these two parts in order to mitigate the risk of exposure to *E. coli* organisms in ground beef servings consumed in the United States.

# 11 CONCLUSIONS BASED ON THE ANALYSES IN THE *E. COLI* MODEL

In this chapter the results of sensitivity analyses with different methods applied to each of the modules and parts of the *E. coli* model are summarized in order to have a better evaluation and comparison of methods. Moreover, the questions raised in the case scenarios in Sections 3.3.1 to 3.3.3 are addressed based on the results of the analyses in Chapters 4 to 10.

This chapter contains two sections. Section 11.1 presents estimated ranks of inputs to the production, slaughter and preparation modules. Section 11.2 addresses questions raised in the case scenario defined for each module.

## 11.1 Relative Rankings of Inputs to the *E. coli* Model Based on Different Sensitivity Analysis Methods

In Chapters 4 through 9, the following methods for sensitivity analysis were applied to different modules and parts of the *E. coli* model:

Chapter 4: Nominal range sensitivity analysis

Chapter 5: ANOVA

Chapter 6: Regression analysis, Pearson (sample) and Spearman (rank) correlation coefficients, and rank regression

Chapter 7: Classification and regression tree

Chapter 8: Scatter plots

Chapter 9: Conditional sensitivity analysis

In Chapter 4, a case study was provided to justify that nominal range sensitivity analysis method was not applicable to the *E. coli* model. Of the other sensitivity analysis methods, six methods produced a numerical ranking of the key inputs either directly or based upon clearly explainable interpretations of results. These methods include ANOVA, standardized linear regression analysis, Pearson and Spearman correlation coefficients, rank regression, and CART. The other two methods, scatter plots and conditional sensitivity analysis, do not provide a clear basis for numerically ranking the key inputs. However, these methods provide insight regarding non-linearities in the model response, thresholds, and interactions based upon which the results from the first three methods can either be verified or refuted. In this section rankings based on the analysis methods providing numerical rankings are compared in order to see how different methods of analysis affect the rank of each input. This comparison enables the evaluation of sensitivity analysis methods and gives insight regarding the unambiguity of rank for each input.

Moreover, because each sensitivity analysis method has specific assumptions regarding functional relationships between the output and inputs, an objective of this section is to ascertain how those assumptions can affect the rank of inputs.

The following three subsections present the comparison of rankings based on different sensitivity analysis methods in the production, slaughter, and preparation modules, respectively.

### 11.1.1 Comparison of Rankings for Inputs to the Production Module

Five sensitivity analysis methods were applied to the production module. These methods include ANOVA, regression analysis, CART, scatter plots, and conditional sensitivity analysis. The production module includes four parts: (1) feedlot prevalence; (2) within feedlot prevalence; (3) breeding herd prevalence; and (4) within breeding herd prevalence. Of the five sensitivity analysis methods, ANOVA, standardized linear regression analysis, and CART produced a numerical ranking of the key inputs either directly or based upon clearly explainable interpretations of results. Scatter plots and conditional sensitivity analysis did not provide a clear basis for numerically ranking the key inputs. However, these methods provided insight regarding non-linearities in the model response, thresholds, and interactions based upon which the results from the first three methods can either be verified or refuted. A comparison of results from different methods is discussed for each of the four parts of the production module. Based upon these comparisons, conclusions are made regarding the appropriateness of each of the sensitivity analysis methods.

A comparison of rankings for sensitive inputs based upon ANOVA, standardized linear regression analysis, and CART is given in Table 11-1 for the feedlot prevalence part. For standardized linear regression analysis, two sets of results are shown. For one set, F values were used to develop the rankings for both qualitative and quantitative inputs, while for the other set, standardized regression coefficients were the basis of the rankings for quantitative inputs. The results from the three sensitivity analysis methods are comparable with respect to identification of the most important input. The fact that three methods with different theoretical underpinnings led to selection of the same input as the top-ranked one suggests that the top ranking is unambiguous and that all three methods responded in a similar and appropriate manner to the importance of the input.

Table 11-1. Comparison of Rankings for Inputs Based on ANOVA, Regression Analysis and CART in the Feedlot Prevalence Part

| Variable | ANOVA | Regression [a] | | CART |
|---|---|---|---|---|
| | | F Value | Coefficient | |
| Study | 1 | 1 | NA [b] | 1 |
| Apparent Prevalence | 3 | 3 | 2 | NS[c] |
| Herd Sensitivity | 2 | 2 | 1 | 2 |

(a) Ranking based on the F Values and the coefficient estimates.
(b) No rank is estimated in regression analysis for qualitative inputs using coefficient estimates.
(c) Not significant based on the analysis.

There is some disagreement regarding the assignment of the second and third ranks to the apparent prevalence and the herd sensitivity. ANOVA and standardized linear regression analysis interpreted based upon F values, and CART all implied the same ranking of these two inputs. The results for standardized linear regression analysis interpreted based upon standardized regression coefficients implied that herd sensitivity was more important than apparent prevalence, which is consistent with the results from ANOVA and regression analysis based upon interpretation of F values. Because the regression coefficient approach could not address the study input, it is expected that the herd sensitivity, if truly of second importance, would be ranked first by this approach. CART did not select apparent prevalence as an input; therefore, this could imply that the output is not sensitive to this input and regression analysis based upon interpretation of F values. Because the regression coefficient approach could not address the study input, it is expected that the herd sensitivity, if truly of second importance, would be ranked first by this approach. CART did not select apparent prevalence as input; therefore, it could imply that the output is not sensitive to this input

The scatter plots presented in Section 8.1.1 illustrated that the model response is nonlinear. Of the three methods compared in Table 11-1, only ANOVA and CART did not impose any assumption regarding model form. The standardized linear regression analysis was based upon a linear assumption. Therefore, of these three methods, the linear regression analysis is expected to have the greatest potential for mis-specification of ranks. However, the ranks from regression were consistent with that of the other two methods, suggesting that the results of regression analysis are unambiguous to departures from linearity in this case.

The comparison of rankings of key inputs for the within feedlot prevalence part is given in Table 11-2. The results from standardized linear regression analysis and CART consistently

Table 11-2. Comparison of Rankings for Inputs Based on ANOVA, Regression Analysis and CART in the Within Feedlot Prevalence Part

| Variable | ANOVA | Regression [a] | | CART |
|---|---|---|---|---|
| Study | 1 | 3 | NA [b] | NS [c] |
| Season | 3 | 4 | NA [b] | NS [c] |
| Apparent within feedlot prevalence | 2 | 1 | 1 | 1 |
| Test Sensitivity | 4 | 2 | 2 | NS [c] |

(a) Ranking based on the F Values and the coefficient estimates.
(b) No rank is estimated in regression analysis for qualitative inputs using coefficient estimates.
(c) Not significant based on the analysis.

implied that the apparent within feedlot prevalence was the most important input. The results from CART implied that this input was substantially more important than any other input, since no other input was included in the regression tree. Moreover, standardized linear regression analysis also indicated that the apparent within feedlot prevalence was substantially more important than other inputs. For example, the regression coefficient of this input differed from the regression coefficient of the test sensitivity by a ratio of approximately seven. The results from ANOVA were qualitatively different than those of the other two methods in that the study was identified as the most important input and the apparent within feedlot prevalence was ranked second. These two ranks were substantially different from each other. The F value for the study differed by a ratio of approximately 12 from the F value of the apparent within feedlot prevalence. In contrast, inputs ranked other than one had comparable importance, because their F values differed by ratios between 1.07 and 1.25.

A comparison of rankings for sensitivity inputs based upon ANOVA, standardized linear regression analysis, and CART is given in Table 11-3 for the breeding herd prevalence part. For standardized linear regression analysis, rankings based upon both F values and regression coefficients are presented. The results from the three sensitivity analysis methods are comparable with respect to identification of the most important input. The top ranking is unambiguous, because all three methods with different theoretical assumptions selected the study as the most important input.

The study was substantially more important than other inputs based upon ANOVA and standardized linear regression analysis when using F values for ranking. The F value for the study differed by ratios of 9 and 4.3 from the F value of the second ranked input when using ANOVA and regression analysis, respectively.

Table 11-3. Comparison of Rankings for Inputs Based on ANOVA, Regression Analysis and CART in the Breeding Herd Prevalence Part

| Variable | ANOVA | Regression [a] | | CART |
|---|---|---|---|---|
| Study | 1 | 1 | NA [b] | 1 |
| Apparent Prevalence | 3 | NS [c] | NS [c] | 3 |
| Herd Sensitivity | 2 | 2 | 1 | 2 |

(a) Ranking based on the F Values and the coefficient estimates.
(b) No rank is estimated in regression analysis for qualitative inputs using coefficient estimates.
(c) Not significant based on the analysis.

Table 11-4. Comparison of Rankings for Inputs Based on ANOVA, Regression Analysis and CART in the Within Breeding Herd Prevalence Part

| Variable | ANOVA | Regression [a] | | CART |
|---|---|---|---|---|
| Study | 1 | 2 | NA [b] | 2 |
| Season | 4 | NS [c] | NA [b] | 4 |
| Apparent within breeding herd prevalence | 2 | 1 | 1 | 3 |
| Test Sensitivity | 3 | 3 | 2 | 1 |

(a) Ranking based on the F Values and the coefficient estimates.
(b) No rank is estimated in regression analysis for qualitative inputs using coefficient estimates.
(c) Not significant based on the analysis.

Of the three methods compared in Table 11-3, only ANOVA and CART did not impose any assumption regarding model form. The standardized linear regression analysis was based upon a linear assumption. Moreover, $R^2$ for the regression analysis was 0.90 indicating that the linear assumption for the model explained 90 percent of the output variability. Therefore, the linear regression analysis results are expected to be reasonably unambiguous.

The comparison of rankings of key inputs for the within breeding herd prevalence part is given in Table 11-4. There is some disagreement regarding assignment of the first ranked input. CART presented substantially a different ranking for inputs, while ANOVA and regression analysis provided approximately comparable rankings. The order of ranks for the top two inputs reversed for the ANOVA and regression analysis. The magnitude of the F values for the linear regression coefficient indicated that the top three inputs were of comparable importance. In contrast, the F values in ANOVA implied that the ranks for the top three inputs were unambiguous, because the F values for those inputs differed substantially. For regression analysis, the $R^2$ of 0.84 implied that the linear relationship assumption between the output and inputs was reasonable.

In the production module ANOVA seemed to present more unambiguous rankings. The results based on ANOVA mostly were supported by other two methods. In addition, high values of $R^2$ implied that rankings based upon regression analysis could be reliable. Graphical methods of sensitivity analysis did not contribute substantially to clarification of the results gained with other sensitivity analysis methods.

### 11.1.2 Comparison of Rankings for Inputs in the Slaughter Module

Three different types of probabilistic analysis were performed for the slaughter module, as described in Section 3.3.2: (1) one-dimensional simulation of variability based upon mean values of uncertain inputs; (2) two-dimensional simulation of variability for each realization of uncertainty; and (3) one-dimensional simulation of both variability and uncertainty co-mingled.

Table 11-5 summarizes ranks of different inputs based on ANOVA, standardized linear regression analysis, and CART for the first probabilistic approach. For regression analysis two sets of ranking are given. The first set presents ranking of inputs based on the magnitude of F values, while the second set gives the rank based on the magnitude of the standardized regression coefficients.

Comparison of rankings in Table 11-5 indicates that the key similarity among the three sensitivity analysis methods is with respect to the identification of the chilling effect as the most important input. ANOVA and regression analysis approximately identified the same inputs in the group of secondary importance inputs, ranked between two and four. In regression analysis although the rank for the most important input was unambiguous, other inputs in the group of secondary importance inputs presented comparable importance. In ANOVA the rank for the chilling effect was ambiguous, as the F value for this input did not differ substantially from the F value of the next important input. A group of secondary importance inputs did not present unambiguous ranking based on the magnitude of their F values.

CART only selected two inputs in the regression tree. These inputs were also selected in the group of top four inputs based upon ANOVA and regression analysis. Regarding statistically insignificant inputs, all three methods identified approximately the same inputs. There was one exception in the ANOVA results, since in this case the number of positive cases at evisceration was grouped with inputs of minor importance.

The $R^2$ for the standardized linear regression analysis was small, indicating that the linear relationship between the output and the inputs was not valid. Therefore, of these three methods,

Table 11-5. Comparison of Rankings for Inputs, Based on ANOVA, Regression Analysis and CART in the Slaughter Module for Variability only Analysis

| Variable | ANOVA | Regression [a] | | CART |
|---|---|---|---|---|
| | | F Value | Coefficient | |
| Total Number of Combo Bin for Each Carcass | NS [b] | NS [b] | NS [b] | NS [b] |
| Total Number of Infected Animals | NS [b] | NS [b] | NS [b] | NS [b] |
| Total Number of Contaminated Animals | 9 | 7 | 4 | NS [b] |
| Probability of Positive Cases at both Steps of Dehiding and Evisceration | NS [b] | NS [b] | NS [b] | NS [b] |
| Number of Positive Cases at both Steps of Dehiding and Evisceration | 5 | 6 | 6 | NS [b] |
| Number of Positive Cases at Evisceration | 8 | NS [b] | NS [b] | NS [b] |
| Chilling Effect | 1 | 1 | 1 | 1 |
| Number of Organisms | 3 | 3 | 4 | 2 |
| Trim Vacuum Washing Efficiency | 2 | 5 | 5 | NS [b] |
| Evisceration Organisms Added | 6 | 8 | 7 | NS [b] |
| Washing Effect | 4 | 2 | 3 | NS [b] |
| Contaminated $cm^2$ | 7 | 4 | 2 | NS [b] |

(a) Ranking based on the F Values and the coefficient estimates.
(b) Not significant based on the analysis.

the linear regression analysis is expected to have the greatest potential for mis-specification of ranks. The fact that three methods with different theoretical underpinnings lead to selection of the same input as the top-ranked one suggests that the top ranking is unambiguous and that all three methods responded in a similar and appropriate manner to the importance of the input. Moreover, results in Table 11-5 indicate that rankings based on the standardized linear regression analysis are substantially comparable to that of the other methods with respect to the selection of key inputs. Thus, even though the $R^2$ value in this case is low, the ranking of the inputs is similar to that obtained with other methods.

ANOVA identified that there were interaction effects between specific inputs. Conditional sensitivity analysis supported the general idea that there were interactions between inputs, but this method did not specifically identify those inputs that have interactions. CART also identified an interaction between the chilling effect and the number of organisms. This interaction was not addressed with ANOVA.

Because they can be automated, ANOVA and standardized linear regression analysis were the only methods considered for the two-dimensional probabilistic approach. Mean ranks based on these two methods in 100 uncertainty realizations are presented in Table 11.6. The key

Table 11-6. Comparison of Rankings for Inputs Based on ANOVA and Regression Analysis in the Slaughter Module for Variability Analysis at Different Uncertainty Realizations

| Variable | ANOVA [(a)] | Regression [(a)] |
|---|---|---|
| Total Number of Combo Bin for Each Carcass | 9.6 | 10.6 |
| Total Number of Infected Animals | 9.4 | 8.3 |
| Total Number of Contaminated Animals | 5.8 | 4.5 |
| Probability of Positive Cases at both Steps of Dehiding and Evisceration | 9.1 | 9.7 |
| Number of Positive Cases at both Steps of Dehiding and Evisceration | 7.5 | 8.1 |
| Number of Positive Cases at Evisceration | 7.4 | 6.8 |
| Chilling Effect | 1.7 | 2.2 |
| Number of Organisms | 4.4 | 4.4 |
| Trim Vacuum Washing Efficiency | 4.2 | 6.3 |
| Evisceration Organisms Added | 8.0 | 6.5 |
| Washing Effect | 4.4 | 6.2 |
| Contaminated $cm^2$ | 5.8 | 4.3 |

(a) Mean ranks in 100 uncertainty realizations.

similarities between ANOVA and regression are with respect to the identification of the most important input, and the group of least importance inputs ranked between 8 and 11. There are some exceptions in inputs identified in the group of secondary importance inputs. Moreover, the frequency of specifying each input as statistically significant in both methods was approximately similar.

Table 11-7 presents ranking based on ANOVA, regression analysis, and CART analysis for the third probabilistic approach. Two inputs were directly ranked by CART analysis (i.e. chilling effect and number of organisms). For ranking other inputs based on CART, regression analysis was used as a complementary sensitivity analysis method.

In ANOVA, ranking for the first important input was unambiguous with substantially different F value from the next important input. The F values for inputs ranked between second and fourth did not differ substantially, indicating that these inputs presented comparable importance. For regression analysis, ranks for the top four inputs were unambiguous with no overlap in the confidence intervals for the regression coefficients.

The $R^2$ for regression analysis was small, indicating that the linear relationship between the output and the inputs was not valid. Therefore, of these three methods, the linear regression analysis is expected to have the greatest potential for mis-specification of ranks. The fact that three methods with different theoretical underpinnings lead to selection of

Table 11-7. Comparison of Rankings for Inputs, Based on ANOVA, Regression Analysis and CART in the Slaughter Module for One-Dimensional Variability and Uncertainty Analysis

| Variable | ANOVA | Regression [a] | | CART |
|---|---|---|---|---|
| | | F Value | Coefficient | |
| Total Number of Combo Bin for Each Carcass | NS [b] | NS [b] | NS [b] | 11 [c] |
| Total Number of Infected Animals | NS [b] | NS [b] | NS [b] | 7 [c] |
| Total Number of Contaminated Animals | 6 | 6 | 4 | 9 [c] |
| Probability of Positive Cases at both Steps of Dehiding and Evisceration | NS [b] | 8 | 6 | 8 [c] |
| Number of Positive Cases at both Steps of Dehiding and Evisceration | 8 | NS [b] | NS [b] | 10 [c] |
| Number of Positive Cases at Evisceration | 7 | NS [b] | NS [b] | NA [c] |
| Chilling Effect | 1 | 1 | 1 | 1 |
| Number of Organisms | 2 | 2 | 2 | 2 |
| Trim Vacuum Washing Efficiency | 3 | 5 | 5 | 3 [c] |
| Evisceration Organisms Added | 5 | 4 | 4 | 5 [c] |
| Washing Effect | 4 | 3 | 3 | 4 [c] |
| Contaminated $cm^2$ | 8 | 7 | 6 | 6 [c] |

(a) Ranking based on the F Values and the coefficient estimates.
(b) NS = Not significant based on the analysis.
(c) Ranked based on a complement sensitivity analysis.

the same input as the top-ranked one suggests that the top ranking is unambiguous and that all three methods responded in a similar and appropriate manner to the importance of the input.

In the slaughter module ANOVA seemed to present more unambiguous rankings. The results based on ANOVA mostly were supported by other two methods. ANOVA also addressed the interactions between inputs. In addition, low values of $R^2$ implied that rankings based upon standardized linear regression analysis might be unreliable, although the similarity between different methods in selecting group of inputs with high importance indicated that the top rankings was robust to the functional assumption of the model. Graphical methods of sensitivity analysis helped in order to identify the interaction effect between inputs.

### 11.1.3 Comparison of Rankings for Inputs to the Preparation Module

In the preparation module sensitivity analysis methods were applied to three parts including growth estimation, cooking effect, and serving contamination parts. This section contains three subsections corresponding to each of these three parts, respectively.

Table 11-8. Comparison of Rankings for Inputs Based on ANOVA, Regression Analysis and CART in the Growth Estimation Part for Variability Analysis at Mean Uncertainty

| Variable | ANOVA [f] | | Regression [a] | | CART [e] | |
|---|---|---|---|---|---|---|
| | Point Estimate | Mean Rank | F Value | Coefficient | Visual Index | Deviance Index |
| Storage Temperature, Stage 1 | 3 | 4.0 | 4 | 3 | 3[c] | 5 |
| Storage Temperature, Stage 2 | 10 | 10.9 | NS [b] | NS [b] | NS[b],[c] | NS[b] |
| Storage Temperature, Stage 3 | 2 | 1.0 | 1 | 1 | 1 | 1 |
| Storage Time, Stage 1 | 4 | 2.9 | 3 | 4 | 4[c] | 3 |
| Storage Time, Stage 2 | NS [b] | 12.3 | NS [b] | NS [b] | NS[b],[c] | NS[b] |
| Storage Time, Stage 3 | 1 | 2.1 | 2 | 2 | 2 | 2 |
| Maximum Density | 9 | 8.1 | 7 | 7 | 9[c] | 8 |
| Lag Period, Stage 1 | 5 | 5.9 | 8 | 7 | 8[c] | 4 |
| Lag Period, Stage 2 | 11 | 10.6 | NS [b] | NS [b] | NS[b],[c] | NS[b] |
| Lag Period, Stage 3 | 6 | 5.1 | NS [b] | NS [b] | 7[c] | 6 |
| Generation Time, Stage 1 | 8 | 7.2 | 6 | 6 | 5[c] | NS[b] |
| Generation Time, Stage 2 | NS [b] | 9.9 | NS [b] | NS [b] | NS[b],[c] | NS[b] |
| Generation Time, Stage 3 | 7 | 6.8 | 5 | 5 | 6[c] | 7 |

(a) Ranking based on the F Values and the coefficient estimates.
(b) NS = Not significant based on the analysis.
(c) Ranked based on a complementary sensitivity analysis.
(d) NA = Rank cannot be evaluated in the complementary analysis
(e) For CART two sensitivity indices are used
(f) For ANOVA two sets of ranking are presented: (1) ranking based on the point estimates of F values; and (2) mean rankings based on the 200 bootstrap simulations

#### 11.1.3.1 Comparison of Rankings for Inputs to the Growth Estimation Part

Three different types of probabilistic analysis were performed for the growth estimation part: (1) one-dimensional simulation of variability based upon mean values of uncertain inputs; (2) two-dimensional simulation of variability for each realization of uncertainty; and (3) one-dimensional simulation of both variability and uncertainty co-mingled.

Table 11-8 summarizes ranks of different inputs based on ANOVA, standardized regression analysis, and CART for the first probabilistic approach. For standardized linear regression analysis, two sets of results are shown. For one set, F values were used to develop the rankings, while for the other set, standardized regression coefficients were the basis of the rankings. For CART, results are presented based on two sensitivity indices, including visualization of the regression tree accompanied by results from complementary analyses and ranking the inputs based on their contribution to the reduction in the total deviance. For ANOVA, a case study was provided to quantify the uncertainty in the F values using bootstrap

technique. Hence, two sets of rankings are presented in Table 11-8 for ANOVA: (1) rankings based upon the relative magnitudes of the point estimates for the F values; and (2) mean rankings based upon 200 bootstrap simulations of the dataset.

A comparison of rankings in Table 11-8 indicates that the key similarity among the three sensitivity analysis methods is with respect to the identification of the four most important inputs. Regarding the group of secondary important inputs, these methods approximately considered the same inputs in this group. Moreover, inputs related to stage 2 were identified to have no statistically significant effect or were placed in the group of least important inputs by all three methods. Two sensitivity indices used in CART for ranking inputs provided approximately the same ranking with respect to the identification of the insignificant inputs and the most important inputs. Two sets of rankings presented for ANOVA are comparable in many ways. In particular, both analyses produced similar rank ordering for groups of factors. Although the numerical values of the ranks from the variability only simulation often do not agree with the average ranks from the bootstrap simulation, the differences can be attributed to random sampling error and the resulting ambiguity in ranks within groups of factors.

Based on the results presented in Section 5.4.2 for quantifying the uncertainty in F values, for a Monte Carlo simulation sample size of 65,000 for the variability only analysis of the growth estimation part, the range of uncertainty in statistically significant F values that were substantially large was found to be approximately plus or minus 30 percent or less. This implies that the F values should differ by 30 percent or more in order to represent rankings that are clearly different.

For the regression analysis, considering the estimated confidence interval for each regression coefficient as a measure of the ambiguity in rankings, inputs ranked between 1 and 4 had unambiguous rankings with no overlap in their confidence intervals for the regression coefficients.

ANOVA identified that there are statistically significant interactions especially between storage time and temperature at stage 1 and 3. This inference was supported by conditional sensitivity analysis. The conditional sensitivity method illustrated that that there is a nonlinear response to the variation of the storage temperature and time before the saturation points.

For the second probabilistic approach, five sensitivity analysis methods were applied to the dataset. These methods include ANOVA, standardized linear regression, Pearson (sample)

Table 11-9. Comparison of Mean Rankings for Inputs Based on ANOVA, Standardized Linear Regression, Pearson Correlation, Spearman Correlation, and Rank Regression Methods in the Growth Estimation Part for Variability Analysis Under 100 Uncertainty Realizations

| Variable | ANOVA | Regression | Pearson | Spearman | Rank Regression |
|---|---|---|---|---|---|
| Storage Temperature, Stage 1 | 4.0 | 3.1 | 4.3 | 6.6 | 9.3 |
| Storage Temperature, Stage 2 | 9.6 | 8.5 | 10.9 | 10.8 | 8.1 |
| Storage Temperature, Stage 3 | 2.5 | 1.4 | 1.6 | 5.1 | 8.0 |
| Storage Time, Stage 1 | 2.7 | 3.5 | 4.7 | 1.9 | 1.5 |
| Storage Time, Stage 2 | 12.1 | 10.8 | 10.9 | 11.1 | 10.4 |
| Storage Time, Stage 3 | 1.3 | 2.7 | 3.9 | 1.7 | 1.6 |
| Maximum Density | 9.7 | 10.1 | 10.6 | 11.1 | 11.2 |
| Lag Period, Stage 1 | 6.4 | 8.9 | 7.0 | 4.9 | 3.6 |
| Lag Period, Stage 2 | 9.2 | 9.7 | 10.8 | 10.7 | 9.0 |
| Lag Period, Stage 3 | 6 | 9.7 | 4.8 | 3.3 | 3.4 |
| Generation Time, Stage 1 | 9.3 | 7.2 | 7.1 | 7.2 | 9.0 |
| Generation Time, Stage 2 | 9.3 | 9.0 | 10.7 | 10.9 | 8.6 |
| Generation Time, Stage 3 | 8.8 | 6.4 | 3.8 | 5.4 | 7.0 |

correlation coefficients, Spearman (rank) correlation coefficients, and rank regression.

According to the results provided in Table 11.9, ANOVA and the two sample-based methods of standardized linear regression analysis and Pearson correlation coefficients produced approximately similar rankings. The key similarity between these methods is with respect to the identification of the top four inputs that have highest mean ranks. There is also similarity in rankings between the two ranked-based methods of rank regression and Spearman correlation coefficients. Generally, results according to the rank-based techniques for sensitivity analysis are different from those of the methods based on the sample data. This difference is more apparent with respect to the inputs to which the model has higher sensitivity. For example, while storage temperature at stage 3 was identified as the most important input using standardized regression analysis and sample (Pearson) correlation coefficients, and as the second important input using ANOVA, this input was attributed low mean ranks of 5.1 and 8.0 using rank (Spearman) correlation coefficients and rank regression methods, respectively. All methods approximately identified the same inputs that have low or no importance. Inputs associated with stage 2 and maximum density were attributed low mean ranks between 8.1 and 12.1.

Table 11-10. Comparison of Rankings for Inputs, Based on ANOVA, Regression Analysis and CART in the Growth Estimation Part for One-Dimensional Variability and Uncertainty Analysis

| Variable | ANOVA | Regression [a] | | CART [e] | |
|---|---|---|---|---|---|
| | | F Value | Coefficient | Visual Index | Deviance Index |
| Storage Temperature, Stage 1 | 4 | 4 | 3 | 3[c] | 6 |
| Storage Temperature, Stage 2 | NS [b] | 8 | NS [b] | NS[b],[c] | NS[3] |
| Storage Temperature, Stage 3 | 3 | 2 | 1 | 1 | 2 |
| Storage Time, Stage 1 | 2 | 3 | 4 | 4[c] | 3 |
| Storage Time, Stage 2 | NS [b] | NS [b] | NS [b] | NS[b],[c] | NS[3] |
| Storage Time, Stage 3 | 1 | 1 | 2 | 2 | 1 |
| Maximum Density | 9 | 7 | 7 | NA[c] | 9 |
| Lag Period, Stage 1 | 5 | 8 | 7 | NA[c] | 8 |
| Lag Period, Stage 2 | 10 | NS [b] | NS [b] | NS[b],[c] | NS[3] |
| Lag Period, Stage 3 | 6 | NS [b] | NS [b] | NA[c] | 7 |
| Generation Time, Stage 1 | 7 | 6 | 6 | 6[c] | 5 |
| Generation Time, Stage 2 | NS [b] | NS [b] | NS [b] | NS[b],[c] | NS[3] |
| Generation Time, Stage 3 | 8 | 5 | 5 | 5[c] | 4 |

(a) Ranking based on the F Values and the coefficient estimates.
(b) NS = Not significant based on the analysis.
(c) Ranked based on a complementary sensitivity analysis.
(d) NA = Rank cannot be evaluated in the complementary analysis
(e) For CART two sensitivity indices are used

Table 11-10 presents rankings based on ANOVA, regression analysis, and CART for the one-dimensional analysis of co-mingled variability and uncertainty. For CART, two alternative sensitivity indices were implemented for ranking the inputs. In the first approach, two inputs were ranked by CART analysis based on visualization of the regression tree (i.e., storage temperature and storage time at stage 3). To rank other inputs based on CART, regression analysis was used as a complementary sensitivity analysis method. In the second approach, the amount of contribution of each input to the reduction of the total deviance was used. The key similarity between these methods of analysis is with respect to the identification of the top four inputs that have the highest ranks. Moreover, all inputs related to stage 2 were identified as statistically insignificant by all three sensitivity analysis methods except for the lag period at stage 2. In the latter case, a low rank of 10 was attributed to this input based upon ANOVA.

The rank for the most important input based on ANOVA was unambiguous, because the F value related to this input was substantially different from the F values of other inputs.

However, inputs with ranks second and third had comparable importance, because their F values did not differ substantially. In contrast, for regression analysis, inputs ranked between 1 and 4 had unambiguous rankings with no overlap in their confidence intervals for the regression coefficients.

ANOVA identified the saturation points in the growth of *E. coli* organisms, considering the interaction between the storage time and temperature. Moreover, conditional sensitivity analysis identified the threshold in the growth of *E. coli* organisms based upon the variation of the storage time and temperature at stages 1 and 3.

In the growth estimation part, ANOVA produced unambiguous rankings. The results based on ANOVA were mostly supported by the other two methods. ANOVA also addressed the interactions between inputs. In addition, moderate values of $R^2$ around 0.50 implied that rankings based upon regression analysis might be unreliable, although the comparison of the results based upon several sensitivity analysis methods indicated that the top ranking was unambiguous to the specific functional assumption of the model, such as linearity. Graphical methods of sensitivity analysis helped to identify the interaction effect between inputs and also identification of thresholds.

#### 11.1.3.2 Comparison of Rankings for Inputs to the Cooking Effect Part

In Table 11-11 rankings based on ANOVA and CART are summarized for the cooking effect part. As explained in Section 6.3.2, regression analysis was not applied to the cooking effect part. Comparison of the rankings indicates that ANOVA and CART analyses identified the same rankings.

Ranking based upon ANOVA was ambiguous, because the F values for the first and second inputs did not differ substantially. This indicates that the cooking temperature and the precooking treatment have comparable importance. In addition, ANOVA identified that there is an interaction between the cooking temperature and the precooking treatment. This interaction was also identified using scatter plots. CART also demonstrated that there is an interaction between these two inputs.

Results from CART implied that there should be a threshold in the model response to the variation of the cooking temperature. CART selected the value of 58 $^{o}$C for the threshold. A review of scatter plots in Section 8.3.2 suggests the presence of a threshold. For example, below temperature of approximately 50 $^{o}$C there is typically no log reduction in the number of *E. coli*

Table 11-11. Comparison of Ranking for Inputs Based on ANOVA, Regression Analysis and CART in the Cooking Effect Part

| Variable | ANOVA | Regression [a] | | CART |
|---|---|---|---|---|
| Precooking Treatment | 2 | | | 2 |
| Cooking Place | NS[b] | | | NS[b] |
| Cooking Temperature | 1 | | | 1 |

(a) The regression analysis was not applied in this part.
(b) NS = Not significant based on the analysis.

organisms due to cooking. At temperatures of approximately 50 $^{o}$C to 58 $^{o}$C there is little log reduction due to cooking depending on the precooking treatment. Above 58 $^{o}$C, there is more sensitivity of the log reduction in the *E. coli* organisms due to cooking for all the precooking treatments. Therefore, the split point chosen by the CART algorithm corresponds, approximately, to a threshold in the model.

11.1.3.3 Comparison of Rankings for Inputs to the Serving Contamination Part

Five sensitivity analysis methods were applied to the serving contamination part. These methods include ANOVA, regression analysis, CART, scatter plots, and conditional sensitivity analysis. The serving contamination part includes two seasons: (1) summer; and (2) winter.

In Tables 11-12 rankings based on ANOVA, regression analysis, and CART are summarized for the summer session. The key similarity between these methods is with respect to the identification of the statistically insignificant inputs to the serving contamination. ANOVA and regression analysis rankings were unambiguous. In ANOVA, F values for the top two inputs differed substantially, and in regression analysis confidence intervals did not overlap, indicating that there was significant difference between inputs based upon these methods. Moreover, all methods identified that there was no statistically significant influence for inputs such as eating location and consumer age.

ANOVA identified that there are interaction effects between inputs. Graphical methods for sensitivity analysis identified that there is a nonlinear response to the variation of the grinder contamination. CART selected specific values for the grinder contamination as a basis for splitting the dataset. Graphical methods supported that the selected value was a threshold.

In Table 11-13 rankings based on ANOVA, regression analysis, and CART are summarized for the winter session. The key similarity between these methods is with respect to

Table 11-12. Comparison of Ranking for Inputs Based on ANOVA, Regression Analysis and CART in the Serving Contamination Part in Summer

| Variable | ANOVA | Regression [a] | | CART |
|---|---|---|---|---|
| | | F Value | Coefficient | |
| Ground Beef Consumption Type | 3 | NS[c] | NA[b] | NS[c] |
| Eating Location | 4 | NS[c] | NA[b] | NS[c] |
| Consumer Age | NS[c] | NS[c] | NA[b] | NS[c] |
| Serving Size | 1 | 2 | 2 | 2 |
| Grinder Contamination | 2 | 1 | 1 | 1 |

(a) Ranking based on the F Values and the coefficient estimates.
(b) No rank is estimated in regression analysis for qualitative inputs using coefficient estimates.
(c) NS = Not significant based on the analysis.

Table 11-13. Comparison of Ranking for Inputs Based on ANOVA, Regression Analysis and CART in the Serving Contamination Part in Winter

| Variable | ANOVA | Regression [a] | | CART |
|---|---|---|---|---|
| | | F Value | Coefficient | |
| Ground Beef Consumption Type | 3 | 3 | NA[b] | NS[c] |
| Eating Location | NS[c] | NS[c] | NA[b] | NS[c] |
| Consumer Age | NS[c] | NS[c] | NA[b] | NS[c] |
| Serving Size | 1 | 2 | 2 | 2 |
| Grinder Contamination | 2 | 1 | 1 | 1 |

(a) Ranking based on the F Values and the coefficient estimates.
(b) No rank is estimated in regression analysis for qualitative inputs using coefficient estimates.
(c) NS = Not significant based on the analysis.

the identification of the statistically insignificant inputs to the serving contamination. All methods identified that there was no statistically significant influence for inputs such as eating location and consumer age.

ANOVA and regression analysis presented unambiguous ranking for inputs. In ANOVA, F values for the top two inputs differed substantially, and in regression analysis confidence intervals did not overlap, indicating that there was significant difference between inputs based upon these methods.

## 11.2 Evaluation of the Proposed Case Scenarios Based on the Results of the Sensitivity Analyses

In Section 3.3 case scenarios were defined for each module and part of the *E. coli* model. Moreover, a few questions were raised in order to have meaningful outcomes from the sensitivity analysis of the *E. coli* model. This section contains three parts. Sections 11.2.1 to 11.2.3 address

the questions raised in Sections 3.3.1 to 3.3.3 for case scenarios in production, slaughter, and preparation modules, respectively.

### 11.2.1 Evaluation of the Case Scenario in the Production Module

In Section 3.3.1 the case scenario for the production module was explained. Three questions were raised based on the case scenario. Those questions are addressed here considering the results of the sensitivity analyses in Chapters 4 to 9.

*Question 1*: What is the ranking of inputs regarding their influence on the output of interest?

The production module has four parts, including feedlot prevalence, within feedlot prevalence, breeding herd prevalence, and within breeding herd prevalence parts. Based on the sensitivity analyses in the feedlot prevalence part, the study effect is considered as the most important input. The herd sensitivity and the apparent prevalence were ranked as second and third important inputs. In the within feedlot prevalence part, most of the sensitivity analysis methods indicated that the apparent within feedlot prevalence is the most important input. Test sensitivity and the study were ranked as second and third inputs, respectively, while the seasonality effect is placed in the last position. For the breeding herd prevalence, the study effect is identified as the most important input, while herd sensitivity and apparent prevalence were identified as second and third important input. Finally, in the within breeding herd prevalence, the study effect was considered as the most important input and the apparent breeding herd prevalence and the test sensitivity were ranked second and third, respectively. The seasonality effect was ranked as the least important input.

The study effect sensitivity indicates that the differing states of knowledge inferred from different studies are important to the assessment. An implication, therefore, is that it may be worthwhile to devote resources to resolve apparent differences among the studies or to collect more representative data in a future studies.

*Question 2*: Is there any study effect in estimation of the response?

The study effect was selected as the most important input. This indicates that the response in each part was affected by the choice of study. Each study has specific parameters such as number of samples, number of positive samples, and the testing method. The importance of the study implies that it is better to use studies that are more representative of the real feedlot

or herd infection prevalence in the United States. These studies should have enough samples and accurate testing methods.

*Question 3*: Is there any seasonality effect for estimation of average within feedlot or breeding herd prevalence?

Seasonality was considered in the within feedlot and within breeding herd prevalence parts. In both parts, the seasonality was ranked as the least important input. Although the seasonality was identified as the least important input, it still presented a statistically significant effect. Thus, seasonality has an effect but it is less important than other inputs.

### 11.2.2 Evaluation of the Case Scenario in the Slaughter Module

In Section 3.3.2 the case scenario for the slaughter module was explained. Four questions were raised based on the case scenario. Those questions are addressed here considering the results of the sensitivity analyses in Chapters 4 to 9.

*Question 1*: What is the ranking of inputs regarding their influence on the output of interest?

The sensitivity analysis methods applied to the slaughter module in Chapters 4 to 9 identified the chilling effect, number of organisms, washing effect, and Trim/Vacuum/Wash efficiency as the top most important inputs. All of the sensitivity analysis methods had agreement on the first important input identified as the chilling effect. Selection of the chilling effect as the most important input by all the sensitivity analysis methods gives insight to the risk managers that careful control of the chilling process is perhaps the most fruitful approach to reduce exposure and/or risk.

*Question 2*: How unambiguous is the identification of key inputs for situations in which variability and uncertainty can be distinguished?

The slaughter module has a two-dimensional variability and uncertainty simulation. Three different types of probabilistic analysis were performed for this module. The key similarity among the three probabilistic simulations was with respect to the identification of the most important input. When regression analysis was used, the results for the variability only and two-dimensional simulations were approximately similar, but both of these differed from the results of the one-dimensional simulation of co-mingled variability and uncertainty. For ANOVA, all three approaches of analysis yielded similar rankings with respect to the most important input, a group of inputs of secondary importance, a group of inputs with minor importance, and a group

of inputs as unimportant. Thus, comparison of the analysis results from these three approaches indicated that since the results were similar, especially in identifying groups of inputs with similar importance, perhaps its is acceptable to use the simplest analysis which is a one-dimensional analysis co-mingling variability and uncertainty in each input.

*Question* 3: Which step in the slaughter module could end up with high contamination in the combo bins?

Based on the sensitivity analyses in the slaughter module, the chilling effect was identified as the most important input. Chilling the carcasses in slaughter plants can lead to an increase in the number of *E. coli* organisms if the storage time and the temperature are not satisfactorily controlled (FSIS, 2001). Chilling effect was identified as the most important input based upon all sensitivity analysis methods. Hence, inadequate chilling carcasses in slaughter plants could cause a high amount of contamination in combo bins. A risk management implication, therefore, is to carefully control the chilling process of the carcasses in the slaughter plants.

*Question 4*: How can the decontamination steps mitigate the number of *E. coli* organisms in combo bins?

Analyses performed in Section 5.3.3 indicated that there was a statistically significant interaction between the chilling effect and the Trim/Vacuum/Wash efficiency. The results of the analyses in Table 5-15 implied that changing the efficiency of the decontamination step from low to high can only affect the contamination in combo bins when there is more than 2 logs increase in the number of *E. coli* organisms during the chilling process. Otherwise, if the amount of *E. coli* organisms on carcasses does not increase more than a 2 logs during the chilling process, there is no statistically significant difference in the final combo bin contamination when applying different efficiencies in the decontamination step (i.e., Trim/Vacuum/Wash step). Thus, a risk management implication is that if in a slaughter plant there is insufficient control regarding the storage time and the storage temperature during the chilling process, more attention should be paid to the decontamination step. With high efficiency during the decontamination process it is possible to decrease in the contamination of the combo bins by approximately 2.6 logs.

### 11.2.3 Evaluation of the Case Scenario in the Preparation Module

In Section 3.3.3 the case scenario for the preparation module was explained. Six questions were raised based on the case scenario. Those questions are addressed considering the results of the sensitivity analyses in Chapters 4 to 9.

*Question 1*: What is the ranking of the input variables regarding their influence on the output of interest in different parts of the module?

The preparation module includes the growth estimation, the cooking effect, and the serving contamination parts. In the growth estimation part, the storage temperature at home, the storage time at home, the storage temperature at retail and the storage time at retail were identified as the top four inputs. Sensitivity analyses indicated that the transportation stage does not have a significant effect on the growth of the *E. coli* organisms. Importance of the storage conditions at home with respect to time and temperature implies that risk managers could consider providing recommendations regarding the storage conditions to consumers as an effective strategy to control the exposure to *E. coli* organisms.

In the cooking effect part, the cooking temperature was ranked as the most important input, while the precooking treatment was identified as the second important input. This finding implies that providing recommendations to the public indicating the minimum cooking temperature can contribute to the reduction of the risk of illness due to exposure to the *E. coli* organisms.

In the serving contamination part for both the high and low prevalence seasons, the grinder contamination and the serving size were identified as the top two important inputs. This finding implies that controls should be focused on the previous steps of the process of bringing foods from farm-to-table before the consumption step, such as in the slaughter plants, as it is perhaps impractical to present recommendations to the public regarding their serving sizes.

*Question 2*: How unambiguous is the identification of key inputs for situations in which variability and uncertainty can be distinguished?

The growth estimation part has a two-dimensional variability and uncertainty simulation. Three different types of probabilistic analysis were performed for this part. The key similarity among the three probabilistic simulations was with respect to the identification of the most important input. When regression analysis and ANOVA were used, the results for all three approaches of analysis yielded approximately similar rankings with respect to the most important

input, a group of inputs of secondary importance, a group of inputs with minor importance, and a group of inputs as unimportant. Thus, comparison of the analysis results from these three approaches indicated that since the results were similar, especially in identifying groups of inputs with similar importance, perhaps it is acceptable to use the simplest analysis approach which is one-dimensional analysis co-mingling variability and uncertainty in each input. However, the comparison of results among probabilistic simulation methods is likely to be case-specific. Moreover, the use of the simplest approach might be useful to identify priorities for data input to the model, but a two-dimensional approach may be required depending on the policy objectives of the analyses to be performed.

*Question 3*: What is the effect of precooking treatments on the log reduction due to cooking?

In the cooking effect part nine precooking treatments were considered. The precooking treatment was identified to have a statistically significant effect. However, based upon the results of the CART analysis, it appears that the first one to three of the precooking treatments produce different results than the others. For these precooking treatments the log reduction in the number of *E. coli* organisms due to cooking is lower than that for the other treatments. The analyses indicated that when the ground beef servings are stored at 15$^{o}$C for nine hours and then 30$^{o}$C for four hours (i.e. the precooking treatment *I* in Table 3-8), the maximum possible log reduction in the number of organisms would occur due to cooking.

*Question 4*: How does the contamination level differ for different age groups?

Results of the analyses in Section 5.4.3 indicated that the consumer age was not a statistically significant input, but it had a significant interaction with other inputs such as the serving size. Hence, based on the analyses, servings consumed by the people between 25 to 64 years old were expected to have higher contamination because of the larger serving sizes for people in this age group.

*Question 5*: What is the effect of eating location on the possible contamination of a ground beef serving?

Analysis in Section 5.4.3 indicated that during high and low prevalence seasons, ground beef servings away from home are more contaminated than the servings consumed at home. Using trim boxes with lower contamination as a part of the meat trim sources for filling the

grinder loads at home leads to lower contamination level of the ground beef servings consumed at home.

*Question 6*: Does the eating place affect the contamination in different ground beef consumption types?

The analysis in Section 5.4.3 indicated that for hamburger patties, ground beef servings are more contaminated away from home during both high and low prevalence seasons, while for meatballs, servings consumed at home are more contaminated in comparison with servings consumed away from home.

# 12 *LISTERIA MONOCYTOGENES* FOOD SAFETY RISK ASSESSMENT MODEL

This chapter describes the model used to estimate the occurrence of the *Listeria monocytogenes* in ready-to-eat foods (RTE). Section 12.1 gives the background on *Listeria monocyotgenes*. Section 12.2 gives a brief overview of the risk assessment of *Listeria monocytogenes*. Section 12.3 explains the structure of the model. The exposure assessment, hazard characterization, dose response and risk characterization of the *Listeria monocytogenes* model are covered in Sections 12.4 and 12.5. The steps in modeling of exposure and dose response are described in Section 12.6. The case scenarios considered for sensitivity analysis are covered in Section 12.7. The model limitations in performing sensitivity analysis and the modifications made to the model are covered in Section 12.8. Finally, Section 12.9 describes the steps in data generation for the application of sensitivity analysis on *Listeria monocytogenes* model.

## 12.1 Background on *Listeria monocytogenes*

*Listeria monocytogenes* is a bacterium often found in soil and water that can cause serious illness. Illness from eating foods with *Listeria monocytogenes* is called “Listeriosis”. Specific groups of people are considered to be susceptible to Listeriosis. However, pregnant women, newborns, older adults, and people with weakened immune systems caused by cancer treatments, AIDS, diabetes, kidney disease, or other illness, are at risk for becoming seriously ill from eating foods that contain *Listeria monocytogenes*. According to the Centers for Disease Control and Prevention (CDC), foods contaminated with *Listeria monocytogenes* cause approximately 2500 cases of illness, including approximately 500 fatalities in the US each year (CDC, 1999b).

Animals can carry *Listeria monocytogenes* in their intestines without becoming sick. As a result, the bacteria may be spread to meat and dairy products. *Listeria monocytogenes* is killed by cooking or by other heating methods, such as pasteurization, used to produce RTE foods. However, RTE foods can become contaminated after processing within the processing plant or along the route from the plant to the plate. Outbreaks of Listeriosis are associated with RTE foods such as hot dogs, luncheon meats, cold cuts, fermented or dry sausage, and other deli-style meat and poultry. In the home, *Listeria monocytogenes* is destroyed if RTE foods are reheated to steaming hot.

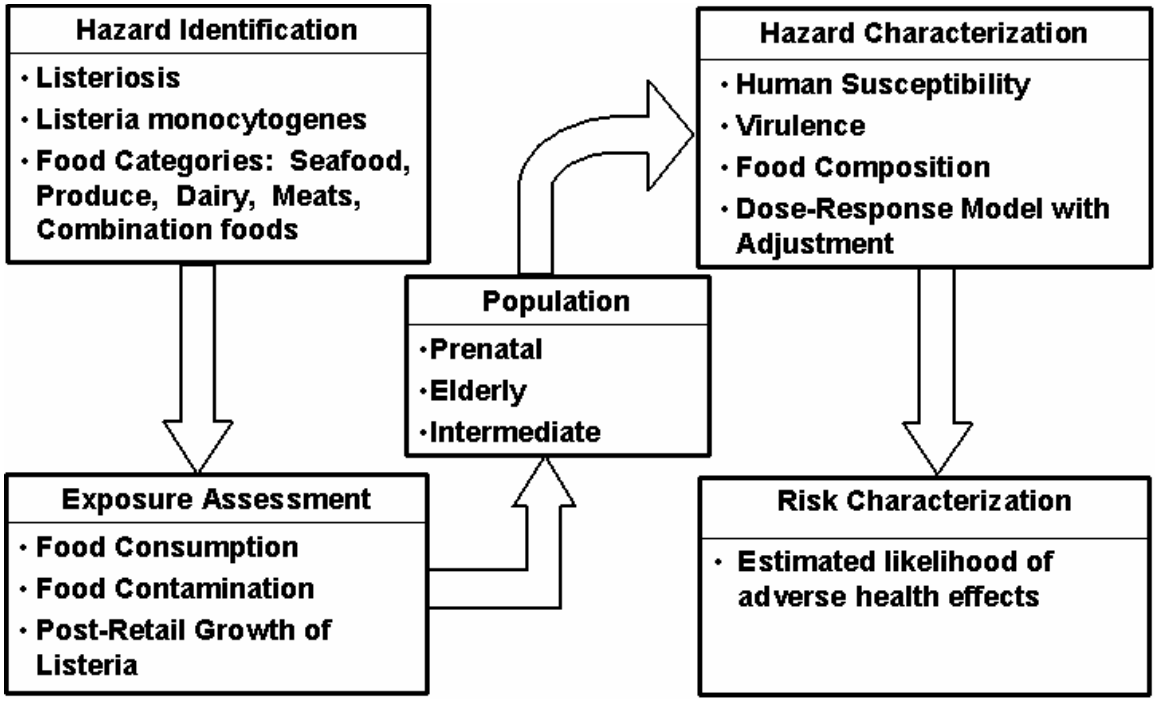



Figure 12-1. *Listeria monocytogenes* in ready-to-eat Food Risk Assessment Model.

## 12.2 Overview of the *Listeria monocytogenes* Model

The U.S. Department of Health and Human Services, Food and Drug Administration's Center for Food Safety and Applied Nutrition (DHHS/FDA/CFSAN) conducted a *Listeria monocytogenes* risk assessment in collaboration with the U.S. Department of Agriculture's Food Safety and Inspection Service (USDA/FSIS) and in consultation with CDC. A food safety process risk model was developed as part of that effort and is referred to here as the "*Listeria monocytogenes.*" The *Listeria monocytogenes* model was used to evaluate the current scientific data and information on Listeriosis (CFSAN, 2001). The model estimated the relationship between exposure to *Listeria monocytogenes* and human susceptibility to illness or death. It followed a framework that separated the assessment activities into four components: hazard identification, exposure assessment, dose-response assessment (hazard characterization), and risk characterization. This framework allowed organization of a highly complex array of varied data, characterization of the predicted consequences, definition of uncertainties, and identification of data gaps. Each component is briefly described and illustrated in Figure 12-1:

- Hazard identification. This part involved the identification of known or potential health effects associated with *Listeria monocytogenes* by establishing the general relationship between the pathogen, its presence in foods, and the adverse outcome (illness or death) associated with consumption of contaminated foods.
- Exposure assessment. This part dealt with the estimation of the likely frequency and level of intake of the pathogen in contaminated foods. It involved evaluation of the probability that the pathogen would be present, the frequency of various levels of contamination, and the impact of food handling, processing, and storage conditions on the overall potential exposure.
- Hazard Characterization. This step dealt with estimation of the relationship between the exposure level (dose) and frequency of illness or other adverse effect (response). The severity of the health effects was also evaluated, often by considering multiple biological endpoints (e.g., infection, morbidity, fatalities).
- Risk characterization. This part consisted of the estimation of the likelihood of an adverse outcome from exposure to the pathogen. The exposure assessment and hazard characterization were combined to mathematically express the probability of adverse effects on given population groups as well as to provide a qualitative or quantitative estimate of the uncertainty associated with the predicted risk values. An important part of this step was determining the degree of uncertainty in relation to the results and distinguishing that from the variation that was inherent in any biological system.

The purposes of the *Listeria monocytogenes* study as summarized by FDA are as follows (CFSAN, 2001):

- To systematically examine the available scientific data and information in order to estimate the relative risks of serious illness and death that may be associated with consumption of different types of RTE foods contaminated with *Listeria monocytogenes*.
- To estimate the potential level of exposure of three age-based population groups of United States consumers to *Listeria monocytogenes* contaminated foods in 20 food categories.
- To relate exposure to public health consequences and estimate the likelihood of human morbidity and mortality.

- To provide a tool for analyzing how to most effectively mitigate the risk of illness from *Listeria monocytogenes* in RTE foods.
- To identify future food-safety research needs.

The next section gives a brief overview of the structure of the *Listeria monocytogenes* model.

## 12.3 Structure of the *Listeria monocytogenes* Model

The *Listeria monocytogenes* model was an extensive study that made a clear distinction between uncertainty and variability dimensions in the exposure assessment. The model was not intensive such as the E. coli model in that it did not get into detailed study for each food category. Instead, the objective was to analyze a large number of food categories using a common framework. The simulation model consisted of exposure and dose response module. The variability and uncertainty in the inputs of the exposure module were propagated to the output by applying two-dimensional Monte Carlo simulations. The dose response module considered only the uncertainty dimension and was modeled using one-dimensional Monte Carlo simulation. The *Listeria monocytogenes* model did not consider seasonality effects in the consumption behavior of the exposed population. Also, the pre-retail food processing steps were not modeled in the *Listeria monocytogenes* model. However, the initial contamination of food in the pre-retail stages was considered as a direct input to the model.

The risk assessment of *Listeria monocytogenes* focused only on severe public health consequences. In general, there were insufficient data to model individual foods. Therefore, 20 food categories as given in Table 12-1 were created in the original study based on: primary origin (seafood, produce, dairy and meat); composition and processing (raw, cooked, pH and salt content); available data on the prevalence of *Listeria monocytogenes* in the foods; and epidemiological information. Consumption data from survey were used for 18 of the 20 food categories (CSFII, 1996). For the other two food categories, data from the NHANNES III study were used (DHHS, 1998). In cases where data was limited or missing, data from similar foods were used (CSFII, 1996 and DHHS, 1998). The latest model considers three additional food categories that are not documented in the *Listeria monocytogenes* draft report. For each food category three sub-populations, including neonatal, intermediate-age and elderly, were separately considered.

Table 12-1. Food Groups Included in the *Listeria monocytogenes* Risk Assessment

| |
|---|
| **SEAFOOD** |
| Smoked Seafood (finfish and mollusks) |
| Raw Seafood (finfish and mollusks) |
| Preserved Fish (dried, pickled, and marinated finfish) |
| Cooked Ready-to-Eat Crustaceans (shrimp and crab) |
| **PRODUCE** |
| Vegetables (raw, dried, and vegetable salads) |
| Fruits (raw, dried, fruit salads, and nuts) |
| **DAIRY** |
| Soft Mold-Ripened and Blue-Veined Cheese |
| Goat, Sheep, and Feta Cheese |
| Fresh Soft Cheese [a] (e.g., queso fresco) |
| Heat-Treated Natural Cheese and Process Cheese (mozzarella, cottage, cream cheese, and cheese spreads) |
| Aged Cheese (hard, semi-hard, and semi-soft cheese) |
| Pasteurized Fluid Milk |
| Unpasteurized Fluid Milk |
| Ice Cream and Frozen Dairy Products |
| Miscellaneous Dairy Products (butter, yogurt, cream) |
| **MEAT** |
| Frankfurters |
| Dry/Semi-Dry Fermented Sausages |
| Deli Meats (cooked, ready-to-eat) |
| Pâté and Meat Spreads |
| **COMBINATION FOODS** |
| Deli Salads (cooked seafood, meat, poultry, egg, and cheese and/or pasta as primary salad ingredients.) |

## 12.4 Exposure Assessment of *Listeria monocytogenes*

An exposure aasessment for foodborn *Listeria monocytogenes* must consider the most significant pathways for exposure and quantitative factors influencing the amount of exposure for any given pathway. The latter addresses the likely consumption levels of the contaminated food. The *Listeria monocytogenes* risk assessment did not consider the contamination pathway or the effects of preventive interventions and controls on the likely consumption levels.

However, the growths during refrigeration and thermal destruction during home cooking or reheating were modeled. Thus, it is possible to gain insight into the importance of consumer controllable actions with respect to exposure.

Exposure is a function of the amount of a food consumed and the level of contamination in that food. Hence, the quantity of contaminated foods likely to be consumed in the U. S. and the levels of *Listeria monocytogenes* in them were estimated in the *Listeria monocytogenes* risk assessment study. Using distributions of contamination and consumption data, estimates of exposure to *Listeria monocytogenes* in the various foods were derived (CSFII, 1996 and DHHS, 1998). Sample weights for weighting the data were used so that they more closely reflect the consumption by the noninstitutionalized U. S. population. The following data were extracted from the food contamination and consumption data:

- The weighted data such as mean amount eaten in grams, median amount eaten in grams and number of servings that characterize all eating occasions in two nonconsecutive days of eating (one day for NHANES III).
- Distributions of the amount of food in grams eaten in all servings over two days for CSFII and one day for NHANES III.
- Distributions of the amount of food in grams eaten in all servings and expressed as weighted percentiles.
- The weighted data values to describe the amount of the food in grams eaten per person per day, as well as the number of eaters.
- The per capita estimates of food eaten.

The data collection for initial food contamination and serving size distribution and annual number of servings are discussed in Section 12.4.1 and 12.4.2, respectively. The growth model to estimate growth between retail and consumption is discussed in Section 12.4.3. The distributions to estimate post retail storage time and maximum growth level for each food category are discussed in Section 12.4.4 and 12.4.5, respectively.

### 12.4.1 Initial Food Contamination and Serving Size Distributions

Contamination data used in the risk assessment were reported from the U.S. and other countries on six continents. Two types of data describing the levels of *Listeria monocytogenes* contamination in food were included in the model:

- Qualitative data for presence/absence such as the number of positive samples relative to the total sample collection.
- Quantitative data such as the number of colony forming units (cfu) that were measured and recorded from the sample. It was conventionally assumed that one cfu is equivalent to one organism.

Qualitative data were converted to an assumed level of 0.04 cfu/g when *Listeria monocytogenes* is present. Thus both qualitative and quantitative data were used in the construction of cumulative distribution curves of *Listeria monocytogenes* levels in food.

Contamination levels at consumption were modeled with the assumption that contamination distributions for a given food in the U.S. do not vary significantly from those in other countries, especially Western Europe where *Listeria monocytogenes* outbreaks had occurred in the past. Similarly, it was assumed that all foods within a category have a similar pattern of contamination. Further more, all *Listeria monocytogenes* food isolates were accepted as having the potential to cause human illness. No differences in ability to grow or other characteristics between food and clinical isolates were assumed. The impact of this assumption should be considered in the uncertainty associated with relative risk determinations.

Three limitations affecting the modeling of the distributions of levels of *Listeria monocytogenes* in foods are discussed here (CFSAN, 2001). First, the occurrence of detectable levels of L. monocytogenes in foods is rare. There are relatively few data points above the limit of detection (0.04 cfu/g). Second, although it was assumed that there is no difference between contamination distributions for foods in the U.S and other countries, the data may not have been representative of food and food processing procedures in the US. Third, there was a wide degree of variation between studies in the occurrence of high levels of *Listeria monocytogenes*. The length of time a food was held at retail before it was obtained for microbial sampling was not recorded in the survey studies. Hence the study assumed that foods were samples without bias and would represent the entire range of post-production and pre-sale conditions for that food.

The frequency distributions of *Listeria monocytogenes* levels at retail in appropriate concentration categories were calculated on one-half logarithmic unit ranges in the original *Listeria monocytogenes* study. The cumulative frequency of occurrence versus the log(cfu/g) was plotted. The resulting data points were fit with parametric models such as Lognormal, Weibull-Gamma, and Beta-Poisson distributions. Because most of the data points were less than 0.04

(cfu/g), the models were fit to the upper tail. The parameter values of the Lognormal, Weibull-Gamma and Beta-Poisson models were optimized using a weighted least squares goodness of fit criterion. The weight accorded to a particular study was proportional to the number of samples in the study. There was no representation of sampling error included in the uncertainty analysis for the distribution of *Listeria monocytogenes*.

The serving size distribution in the *Listeria monocytogenes* model was handled as an empirical distribution to describe the serving sizes in terms of grams of food eaten per servings in the 20 food categories. These distributions were expressed as percentiles of the amount of food eaten per serving, weighted to reflect the consumption survey demographics. Empirical distributions were used for serving sizes. There were no uncertainties assigned for these food categories. The *Listeria monocytogenes* draft lists three reasons as to why uncertainty was not considered related to the serving sizes (USDA/ARS, 1998a, 1998b):

- Even the smallest data sets used to characterize the serving size distributions were large relative to other inputs to the model.
- Although the data was not completely representative of the current population of the U.S, the data came from a survey explicitly designed for that purpose.
- The variability in intake covered a smaller range than many other parts of the model.

Most of the contamination data used in the *Listeria monocytogenes* model was from samples taken during retail or storage prior to retail. For better estimation of the number of *Listeria monocytogenes* organisms consumed for each food category, the possibility of pathogen growth was considered. The next section covers how the calculation of growth between retail and consumption was done in the *Listeria monocytogenes* study.

### 12.4.2 Annual Number of Servings of Foods

In order to estimate the number of servings of food in each food category, two key assumptions were made. First, the numbers of servings in a specific food category were extrapolated from short-term surveys to an annual basis. Second, the numbers of eaters of a food per day were extrapolated to an annual basis from short-term surveys. Some foods are unlikely to be eaten with the same frequency during a long term basis compared to what may be observed in a short term one or two day survey. Furthermore, some foods are not available year round, or people may not purchase more costly items for regular consumption. For example, pâté and smoked seafood are often higher priced delicacy items.

### 12.4.3 Growth Between Retail and Consumption

The *Listeria monocytogenes* study incorporated a growth model in exposure assessment to consider growth between retail and consumption. The growth model included a specific function for the growth rate of this pathogen as given in Equation 12-1 and several inputs, such as initial contamination level for *Listeria monocytogenes* in the food at retail where the food was purchased, the storage temperature in the home refrigerator, the storage time in the house, and the maximum growth. Higher values for refrigeration temperature were assumed typically to lead to faster growth in time, but the storage time and the refrigeration temperature are not independent. Long storage time and high refrigeration temperature were considered improbable to happen simultaneously as this combination would lead to noticeable spoilage of the food, in which case the food would not be consumed. The output from the growth model was a frequency distribution, indicating the contamination level per gram of each food category at the time of consumption.

A square root model for exponential growth rate (EGR) was incorporated into the model as given in Equation 12-1 because of its simplicity and frequent use in the microbiology literature (Ratkowsky et al., 1982).

$$\sqrt{EGR} = a(T - T_0) \qquad (12\text{-}1)$$

where,

| | | |
|---|---|---|
| EGR | = | exponential growth rate ($\log_{10}$ cfu/day) |
| T | = | growth temperature ($^0$C) |
| $T_0$ | = | extrapolated minimum notational growth temperature ($^0$C) |
| a | = | slope parameter for *Listeria monocytogenes* in the specific food |

$T_0$ values were estimated from four sources (Alavi et al., 1999; Duh and Schaffner, 1993; USDA, 1997; Wijtzes et al., 1993) and the average of these values (-1.18°C) was used in the model.

The *Listeria monocytogenes* study used different storage temperatures from the published literature that reported growth of *Listeria monocytogenes* in various foods. As the growth data were at different temperatures they were converted to equivalent EGRs ($\log_{10}$ cfu/day) at 5 $^0$C as

given in Equation 12-2. The equation for the ratio of EGR at 5 $^0$C to EGR at the reported study temperature is:

$$\frac{EGR_5}{EGR_{lit}} = \left[\frac{a(T_5 + 1.18)}{a(T_{lit} + 1.18)}\right]^2 = \left[\frac{6.18}{(T_{lit} + 1.18}\right]^2 \quad (12\text{-}2)$$

where,

| | | |
|---|---|---|
| $EGR_5$ | = | converted growth rate at 5 $^0$C ($\log_{10}$ cfu/day) |
| $EGR_{lit}$ | = | growth rate from the inoculated pack study ($\log_{10}$ cfu/day) |
| $T_5$ | = | set to 5 $^0$C to standardize the EGRs ($^0$C) |
| $T_{lit}$ | = | temperature used in the literature ($^0$C) |

The modeling process used a cumulative distribution of EGR from the data points in the published literatures. If a food category had five or more data points, different statistical distributions were fitted to the cumulative frequency distribution of EGR with the residual sums of squares for each frequency distribution used to weight the distributions. The probability of each growth model dictated the frequency of selection of each distribution for use in uncertainty iterations during a Monte Carlo simulation (Cassin, et al., 1998; Vose, 1998). For food categories with less than five data points, a triangular distribution defined by the minimum, mode, and the maximum values of EGR was used. For the food categories that had two data points, a uniform distribution was used. The list of parameters for models fit to EGR is listed in Appendix 5 of the *Listeria monocytogenes* draft (CFSAN, 2001).

The data for home refrigerator temperatures were obtained from a 1999 survey conducted by Audits International. The total number of samples was 939 refrigerators, and 26% of the refrigerators exceeded 5 $^0$C. The refrigeration temperatures were modeled with an empirical distribution where values were interpolated from the table of frequency for refrigeration temperature ranges provided by Audits International. For the estimation of the amount of *Listeria monocytogenes* growth occurring between retail purchase and the food consumption, storage time and the EGR were multiplied.

### 12.4.4 Post-Retail Storage Time for Food Categories

Some foods are consumed on the day of purchase whereas others remain in the home refrigerator for lengthy periods of time. This was a major source of variability in storage time

considered in the *Listeria monocytogenes* study. However, except for frankfurters and deli meats, no data were found on the storage of foods in the home. Therefore the *Listeria monocytogenes* study used expert judgments of individuals familiar with the production and use of the various foods to estimate storage time, including variation and uncertainty. The variation in storage time was described using a BetaPert distribution. The *Listeria monocytogenes* model assumed a negative correlation between storage temperature and time. Thus, the uncertainties in the most likely and the maximum storage time were negatively correlated to the temperature.

The uncertainty in storage time was described using a –20% to +20% uniform distribution for the most frequent value, and a –50% to 50% uniform distribution for the maximum value, assuming 100% correlation between these two distributions. The *Listeria monocytogenes* food risk assessment model estimated consumer food practices, not necessarily the recommended storage times.

### 12.4.5 Maximum Growth Level

The estimated growth during storage was added to the contamination level at retail for every iteration step of the Monte Carlo sampling. However, the *Listeria monocytogenes* study did not consider a lag phase in growth; hence it was assumed that the *Listeria monocytogenes* cells were already in the food and adjusted to the food's environment during the period before retail purchase. The only change made from retail to storage was to a new refrigerator temperature.

For each food category, a maximum growth level for *Listeria monocytogenes* was considered based upon published literature (Appendix 8, CFSAN, 2001). Thus, the value for *Listeria monocytogenes* concentration estimated by the model was compared with maximum growth possible for the food category and the smaller of the two was selected. The maximum growth levels (cfu/g) used were applied across all food categories with $10^5$, $10^{6.5}$ and $10^8$ used as maximums for temperatures of <5, 5 to 7 and >7°C, respectively. For milk, sufficient data was available to estimate growth levels of $10^7$, $10^{7.5}$ and $10^8$ at the three storage temperatures, respectively.

## 12.5 Hazard Characterization, Dose Response, and Risk Characterization

The exposure assessment of *Listeria monocytogenes* food safety risk assessment was covered in Section 12.4. This section covers the next three stages of food risk assessment

framework. Hazard characterization is described in Section 12.5.1, dose response in Section 12.5.2 and risk characterization in Section 12.5.3.

### 12.5.1 Hazard Characterization

Hazard characterization describes the adverse effects of a particular substance, organism, or other entity. In the case of *Listeria monocytogenes*, the overall incidence of illness, its severity, and the differential risk to immunocompromised subpopulations were well characterized. In contrast, the relationship between the amounts of *Listeria monocytogenes* consumed and the likelihood and severity of illness resulting from that dose were not well understood. The *Listeria monocytogenes* risk assessment focused on characterization of the dose-response relationship.

The study used surveillance data to describe the magnitude and the incidence of severe disease. The dose-response relationship for the intermediate-age subpopulation used human data from surveillance studies and data from surrogate studies using animals. An adjustment factor was applied to the elderly and perinatal subpopulations to account for increased host susceptibility. This adjustment factor used animal data to establish a susceptibility range, and human epidemiological surveillance data to adjust for increased susceptibility of these subpopulations.

The *Listeria monocytogenes* draft risk assessment considered neonatal deaths result from food borne infection of a pregnant woman, which then is transmitted, to the fetus before or during birth. The neonatal death rates were adjusted to include prenatal infections that resulted in very early termination of pregnancy (i.e., miscarriages). Distinct disease surveillance data on prenatal deaths were not consistently reported available and was estimated based on the reporting of Listeriosis infections for the mother. An adjustment is made in the risk characterization section to include all perinatal deaths that consider both prenatal and neonatal deaths.

### 12.5.2 Dose Response

The dose response modeling of *Listeria monocytogenes* included sources of uncertainty such as food matrix, virulence and human susceptibility. An adjustment factor was used to account for sources that were not considered. Susceptibility and virulence data were combined with the mouse studies to generate a dose-response model to predict the percentage of the three age-based subpopulations that would become ill after being exposed to a particular dose. The

dose-response model can thus predict the number of deaths for any level of exposure, but a single exposure level can also be used to compare the three age-based groups.

The dose-response is a function of the number of *Listeria monocytogenes* consumed and their virulence. The *Listeria monocytogenes* draft listed three factors that affect the dose-response relationship:

- Food matrix: The composition of a food, referred to as the food matrix, affects the ability of pathogens to survive inside the body and cause virulence.
- Virulence: Different strains of *Listeria monocytogenes* vary in their ability to cause illness. This variability influences the number of organisms required to produce illness and possibly the severity or symptoms of illness.
- Human susceptibility: Immunological and physiological factors in humans play a role in determining the distribution of susceptibility that may be found throughout a population. The probability of death was described for the three different age-based groups of people.

Because of variability in host susceptibility and food matrix effects, there is no single infectious dose for *Listeria monocytogenes*, or any other pathogen, that can be used for all individuals. The study used surrogate data from animals or artificial environments to derive dose response curves since *Listeria monocytogenes* can be fatal for humans participating in the experiment. The dose response curves for animals were modified for strain variation, host susceptibility, and differences between mice in a controlled lab environment and humans in an uncontrolled natural environment. The dose response curve in the study predicted the morbidity or number of deaths corresponding to a given dose. The variability in host susceptibility and food matrix was taken into account by adjustment factors, as there was not enough information available to model the variability in the process.

A dose-response adjustment factor was applied to the dose-response model to align the range of predicted numbers of deaths with the current epidemiological information. Without the adjustment, when the mouse dose-response model is coupled with the human exposure assessment model, the model can overestimate the incidence of lethal infections in humans from *Listeria monocytogenes* by a factor of over one million. The study attributed this large overestimation to the lower susceptibility of humans compared to laboratory mice. The

adjustment factor accounted for all of the possible known factors, as well as unknown factors, that may influence virulence. The magnitude of the adjustment factor would change if any of the

Table 12-2. Parameters for Variability Distributions for Host Susceptibility for Listeriosis

| Distribution | Minimum | Most Frequent | Maximum |
|---|---|---|---|
| Low Variability | -1 to 0 | 0 | 0 to 1.5 |
| Medium Variability | -1 to 0 | 0 | 1 to 3 |
| High Variability | -1 to 0 | 0 | 2.5 to 4.5 |

currently accounted factors are revised or enhanced. Also if new factors were accounted for in future, the adjustment factor magnitude would change.

The variation in host susceptibility was represented with triangular distributions. In order to represent populations with low, medium and high ranges of susceptibility, three alternative triangular distributions were applied to generate three different effective dose estimates. The distributions had a minimum value of -1 and a mode value of 0, so that the net effect of the host susceptibility adjustment was to broaden the distribution of effective doses without greatly altering the midpoint. The maximum values for the three distributions were 1, 2.5, and 3.5 $\log_{10}$ cfu for the low, medium, and high variability populations, respectively as shown in Table 12-2. In addition, the tails of the frequency distributions of host susceptibility were assigned uncertainty ranges using uniform distributions, so that there was overlap in the uncertainty ranges of the three frequency distributions.

The *Listeria monocytogenes* study used high variability host susceptibility distributions for the intermediate age and elderly sub-populations since the members of these sub-populations were most probable to exceed the range of physiological states characterized by the animal research. Since the susceptibility of the elderly or immuno-compromised individual could vary, wider ranges are assigned to these groups. The prenatal dose-response functions were based on the medium variability distributions since the basis of categorization of population was not based on degree of immunity or susceptibility. The three host susceptibility distributions encompassed the range of susceptibility that was observed in animal studies. In conjunction with a population-specific dose response adjustment factor, these distributions were used to create a unique dose response function for a particular subpopulation.

The neonatal, intermediate and elderly dose-response curves are shown in Figure 12-2 to 12-4, respectively. The figures show the dose required to produce death from a series of servings. The factors that were responsible for uncertainty in curves are: (1) the variability in the virulence of different strains and the uncertainty in the animal data used to characterize those strains;

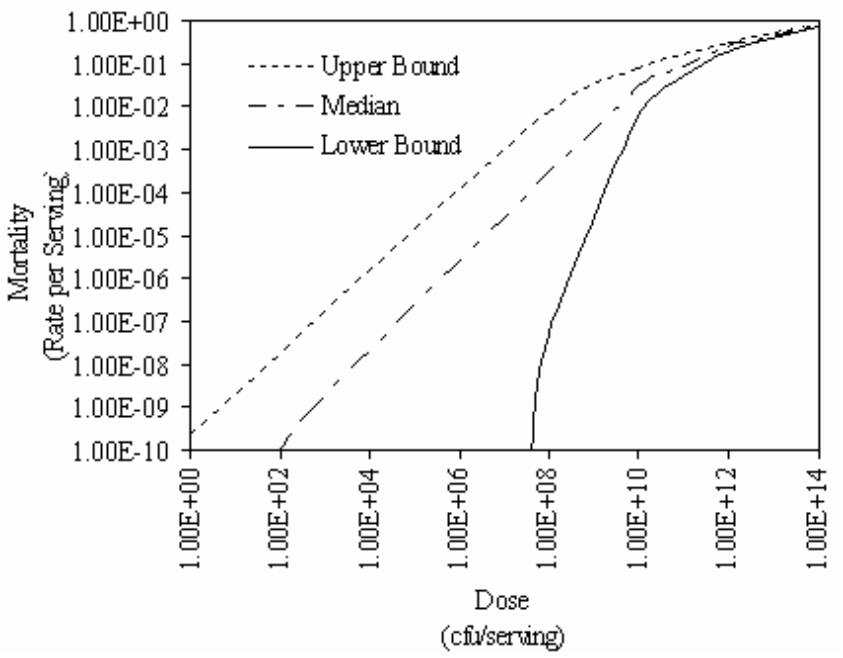



Figure 12-2. *Listeria monocytogenes* Dose-Response Curve with Variable Strain Virulence for the Neonatal Sub-Population (CFSAN, 2001).

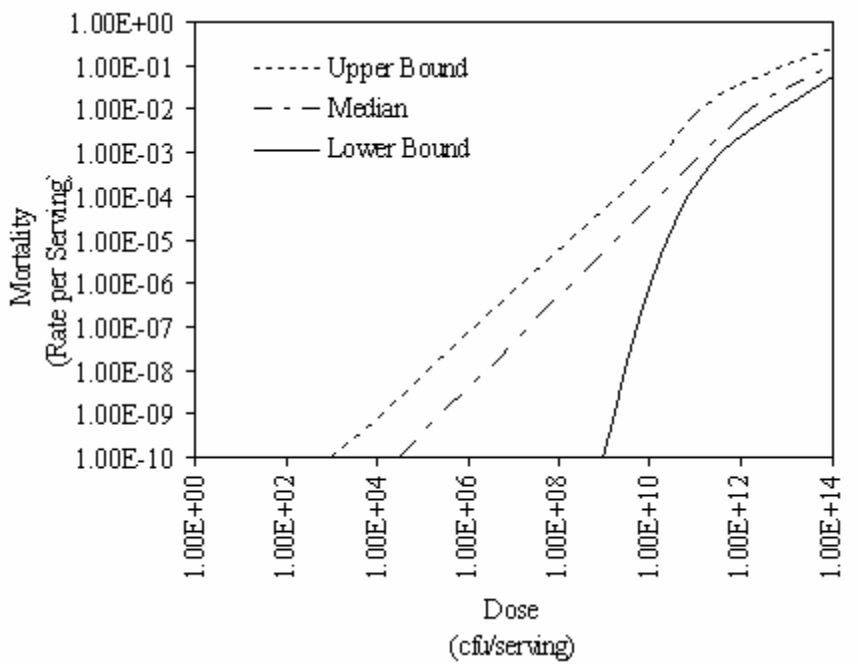



Figure 12-3. *Listeria monocytogenes* Dose-Response Curve with Variable Strain Virulence for the Elderly Sub-Population (CFSAN, 2001).

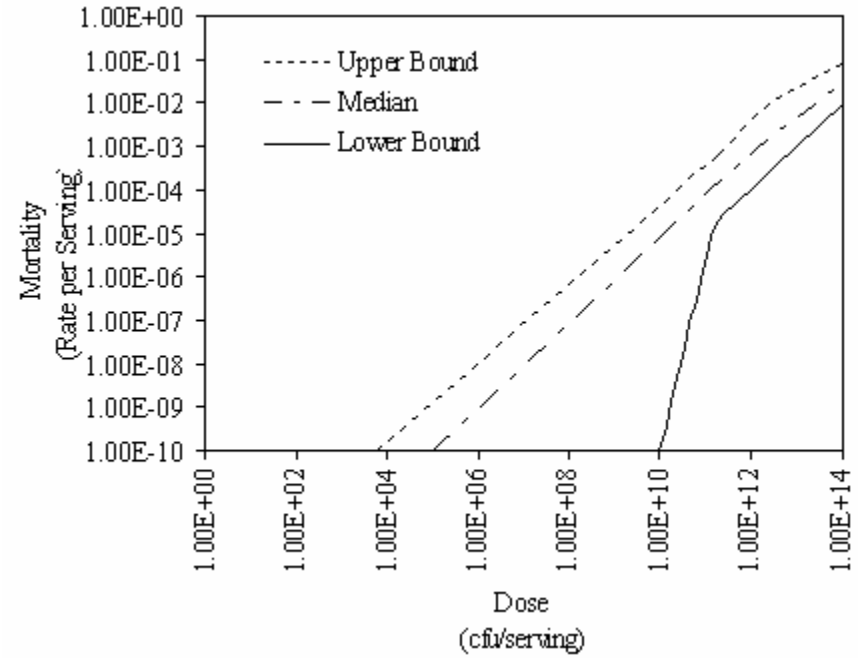



Figure 12-4. *Listeria monocytogenes* Dose-Response Curve with Variable Strain Virulence for the Intermediate-Age Subpopulation (CFSAN, 2001).

(2) the variability in animal susceptibility and the uncertainty in the animal data; (3) the variability and uncertainty in the primary mouse model curve; and (4) the uncertainty in the dose-response adjustment factor. For example, in Figure 12-3 at a dose of $1 \times 10^{10}$ cfu/serving,

the dose-response model predicts a median death rate of 1 in 27,000 servings for the elderly sub-population. However, the uncertainty introduced by the variability in virulence and in host susceptibility provides a lower bound prediction of 1 death in 2 million servings and an upper bound prediction of 1 death in approximately 4,300 servings.

The combined prenatal and neonatal deaths were 2.5 times the neonatal deaths (Buchholz, 2000). The final risk characterization described the perinatal deaths as both prenatal and neonatal.

### 12.5.3 Risk Characterization

In risk characterization, the adverse effects likely to happen in the population are estimated. The probability of contracting *Listeria monocytogenes* from consumption of a single serving of food in one of the 20 food categories was estimated. Risk per annum was estimated based on the annual number of servings.

The *Listeria monocytogenes* study did not consider a dose-response relation for infection or serious illness in risk characterization part. The number of serious illnesses was estimated to be five times the number of deaths based upon 1997 FoodNet data (CDC, 1998a). This factor of five was used in the *Listeria monocytogenes* study to estimate the number of serious illnesses, including deaths, in the risk characterization, as it more accurately reflected the total number of food borne Listeriosis cases.

## 12.6 Modeling of Exposure and Dose Response

This section explains the modeling algorithm used in the *Listeria monocytogenes* food risk assessment model. The steps in modeling are explained first and then the simulation details of the Monte Carlo technique are presented.

Figures 12-5 and 12-6 depict the risk assessment process. The exposure assessment steps are given in Figure 12-5. In Figure 12-6, the hazard characterization steps are in medium gray boxes, and the risk characterization steps are in dark gray boxes. The numbers in the circles indicate the sequence of calculations in the model. The steps shown in the figures are listed below:

Step 1. Distributions for contamination at retail for each food category are assigned.

Step 2. Distributions for the reference growth rate at 5°C for each food category are assigned.

Step 3. A distribution of home refrigerator temperatures in the United States is assumed and is the same for all food categories.

Step 4. Distributions for post-retail storage time for each food category are assigned.

Step 5. A growth model is used for all food categories. This section calculates the exponential growth rate for the specified refrigeration temperature and multiplies by the storage time.

Step 6. The maximum allowable *Listeria monocytogenes* concentration for each food category is checked here. Post growth *Listeria monocytogenes* concentrations are truncated at this level. The maximum growth is temperature dependent with more growth allowed at higher refrigeration temperatures.

Step 7. A model representing the effect of reheating frankfurters on *Listeria monocytogenes* concentration, used for frankfurters only, is considered.

Step 8. Calculates the net contamination at time of consumption using inputs from steps 1, 6, and 7.

Step 9. The distributions of serving size for each food category are assigned at this step.

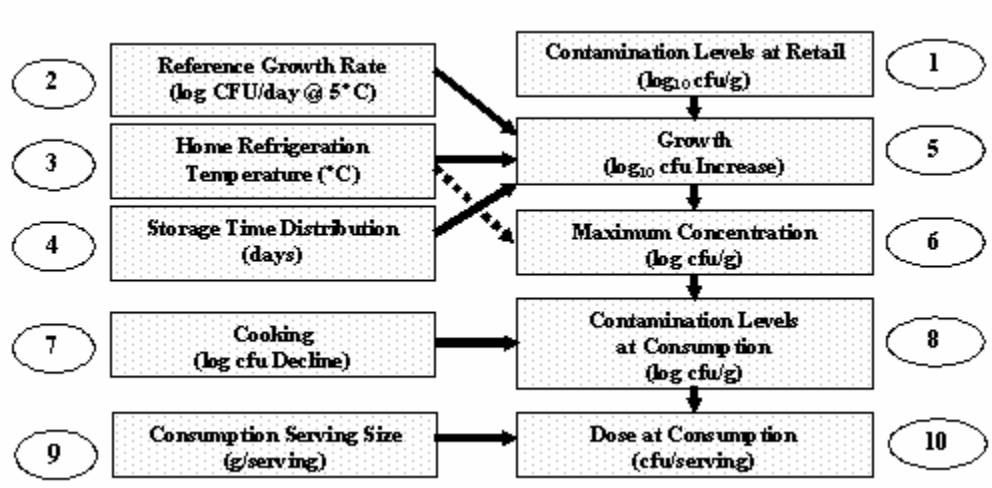



Figure 12-5. Flow Chart of *Listeria monocytogenes* Risk Assessment Model for Individual Exposure Components (CFSAN, 2001).

Step 10. The distributions of dose at consumption for each food category is the final output of the two dimensional simulation. After collapsing the variability dimension to half-log dose bins, the output for each food category is conveyed to the one dimensional dose response simulation for each population group.

Step 11. A distribution for variability of *Listeria monocytogenes* strain virulences in mice is assigned, with the implicit assumption that a similar range will be observed in humans.

Step 12. A distribution adjusting for variability in host susceptibility among humans is assigned, with three (High, Medium, Low) separate adjustments applied to represent different possible ranges. The adjustment increased the range of effective doses.

Step 13. The sum of the strain variability (Step 11) and host susceptibility distributions (Step 12) obtained by two-dimensional Monte-Carlo, with 100,000 variability iterations and 300 uncertainty iterations. The variability dimensions were collapsed to half log dose bins.

Step 14. Summation of the exposure assessment (Step 10) and adjustment factor (Step 13) for each food category is done at this step.

Step 15. The annual number of meals consumed for each food category is calculated.

Step 16. The dose-response adjustment factor is applied in order to make the predictions consistent with CDC estimates of the annual death rate attributable to the population group (i.e., the median value in Step 22).

Step 17. The number of annual servings falling in each dose bin for each food category is calculated here. This is obtained by multiplying the number of servings (Step 15) by the fraction falling in each effective dose bin (Step 14).

Step 18. The death rate per serving for each dose bin (from step 14) is calculated, using the dose-response function derived from mouse data.

Step 19. The intermediate number of annual deaths for each of the dose bin and the food categories is calculated. This was obtained by multiplying the death rate per serving (Step 18) by the number of servings for the dose bin (Step 17).

Step 20. The death rate per serving for each food category was calculated by summing across dose bins. This is obtained by summing the product of the death rate (Step 18) and serving fraction (Step 14) across all bins.

Step 21. The annual number of deaths for each food category was calculated by summing across dose bins (Step 19).

Step 22. The total number of deaths was calculated by summing across food categories.

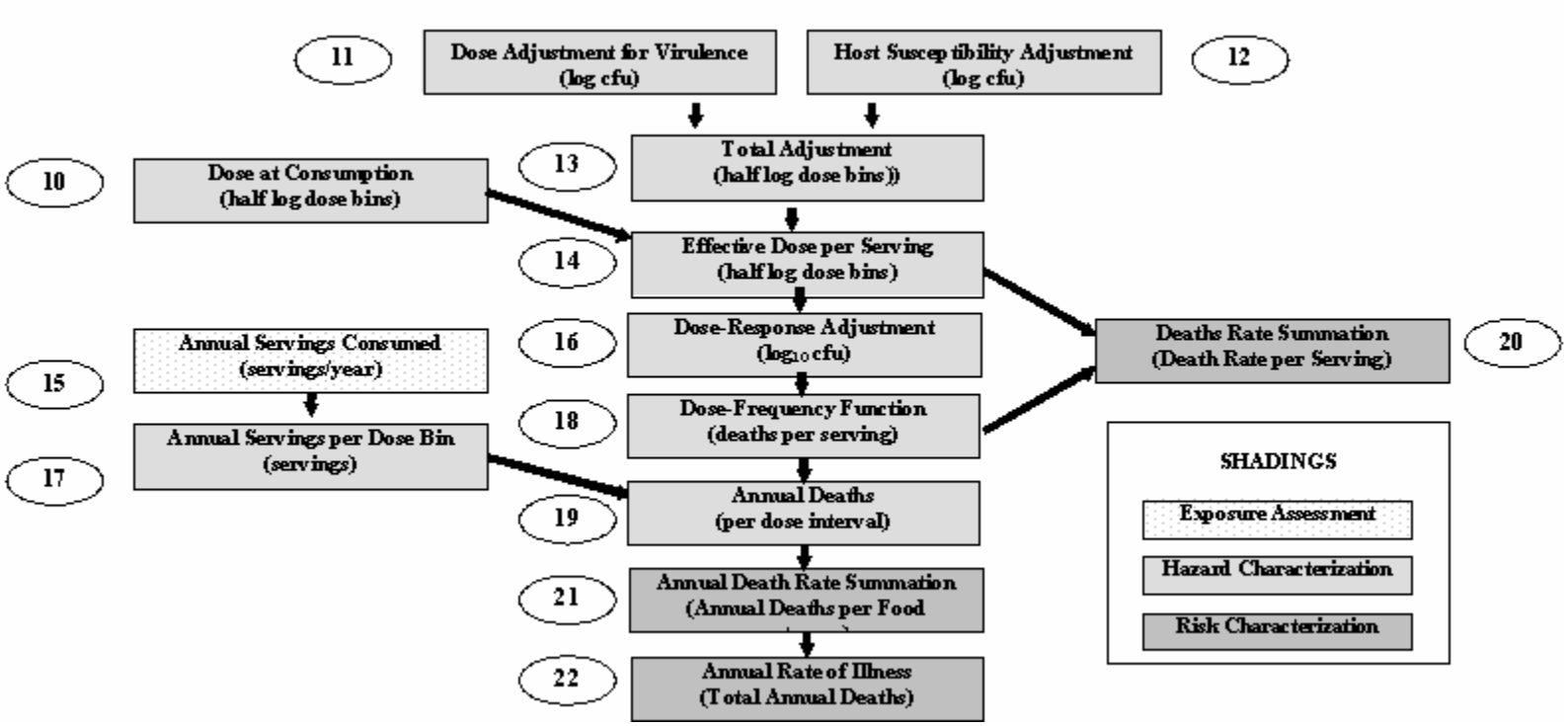



Figure 12-6. Flowchart of *Listeria monocytogenes* Risk Assessment Calculation of Population Estimates (CFSAN, 2001).

To model the rare occurrence of Listeriosis direct application of Monte Carlo modeling did not provide adequate characterization of the tails of the distributions in the model. Therefore, the study divided the model into two major components: (1) the exposure assessment; and (2) the dose response adjustment factors. Each of these components of the model covered 10 to 15 $\log_{10}$ ranges. The simulations in the original study were as follows (CFSAN, 2001):

- A two dimensional Monte Carlo simulation was used in the exposure models for each of the food groups, with 30,000 variability iterations and 300 uncertainty iterations. A common set of random numbers was used to represent variability and uncertainty for all of the twenty-food categories.
- A two dimensional Monte Carlo simulation was used to estimate the variability and uncertainty of the strain virulence and host susceptibility functions, with 100,000 variability iterations and 300 uncertainty iterations.
- The variability dimension for the above two simulation was condensed to 42 half-$\log_{10}$ bins, which ranged from -5 to +10 logs for each of the 300 uncertainty iterations.
- During the one-dimensional uncertainty-only dose-response simulation, dose bins from the exposure assessment for each food group were combined with the strain virulence and host susceptibility dose bins.

The exposure assessment modeled the effect of various factors such as frequency and extent of contamination at retail, consumption patterns, the growth potential of *Listeria monocytogenes* in foods, length of refrigerated storage, and refrigeration temperatures. The dose values that considered both initial *Listeria monocytogenes* concentration at retail and growth between retail and consumption were combined with the three dose-response models for the susceptible subpopulations to yield predictions of the relative role of each of the 20 food categories in Listeriosis in the United States, on a per serving and a per annum basis. The risk characterization was anchored such that the overall predicted incidence of Listeriosis was consistant with the actual incidence of Listeriosis. An implicit assumption was that the foods encompassed by the 20-food categories account for all cases of foodborne Listeriosis. This part of the modeling was done using the Excel “goal seek” function. Changing the adjustment factor for doses attained the target of actual number of known deaths and hence the mortality for changed dose. Thus the goal seek function was used to calculate the adjustment factor. The

relative rank of the medians of the 4,000 uncertainty iterations for each food category and each subpopulation for the per annum predictions are reported in the *Listeria monocytogenes* draft risk assessment.

### 12.7 Case Scenarios

A detailed sensitivity analysis on the *Listeria monocytogenes* model calls for extensive computational resources both in terms of time and space. Case scenarios were defined so as to narrow the scope of analysis. For example, it was deemed more important to focus on selected food categories, based on expert recommendation. Different subpopulations vary in susceptibility and therefore the analysis was concentrated on a specific subpopulation. The methodology of sensitivity analysis is not dependent on these factors that reduce the scope of the problem. The factors simply change the domain of application of the methods. Thus, to demonstrate the methodological aspects of the sensitivity analysis methods a smaller domain was considered.

In consultation with Dr. Peter Cowen and Dr. LeeAnn Jaykus of North Carolina State University and Dr. Clark Clarington of FDA, pâtés and meat spreads, milk, smoked seafood, fresh soft cheese and deli salad food categories were identified as priorities for the analysis. Pâtés and meat spreads were considered in the same food category. Pâtés include hotdogs that are accepted as the most important source of Listeriosis based on survey data (CSFII, 1996 and DHHS, 1998). Although the prevalence of *Listeria monocytogenes* in milk is low, its consumption rate is high and it is estimated to account for a large portion of Listeriosis deaths (CFSAN, 2001). The prevalence rate of smoked seafood is high. The largest outbreak of Listeriosis in US was attributed to fresh soft cheese (CFSAN, 2001). Deli salad was chosen as it has a high potential for contamination due to extensive handling preparation. Neonatal and elderly sub populations have high susceptibility and incidence rate. However only the neonatal sub-population was selected in order to narrow the scope of the study.

The *Listeria monocytogenes* model separates variability from uncertainty. Sensitivity analysis was performed specifically considering only variability and both variability and uncertainty together. Comparisons between these different case scenarios help in understanding the effect of an assumption on ranking of the importance of the input variables. Variability analysis on the exposure module was also performed for different uncertainty realizations. In principle, uncertainty analysis on the exposure module under different variability realizations is possible to perform.

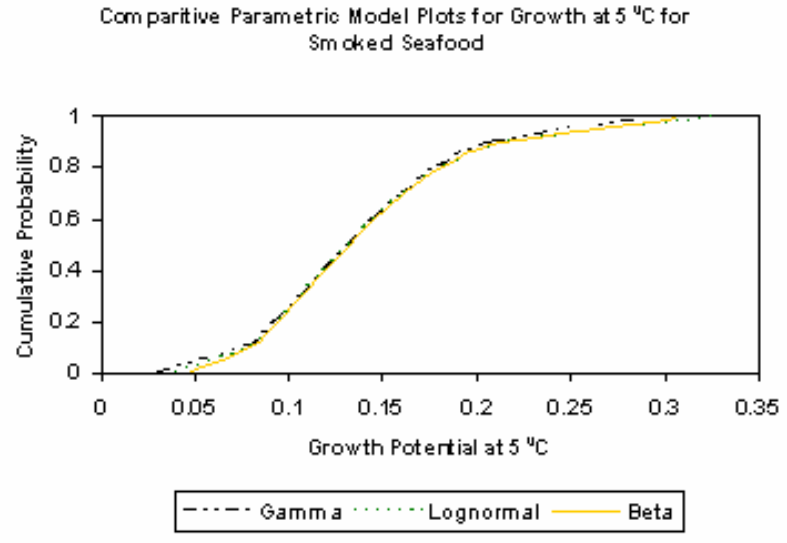



Figure 12-7. Plot of the Three Parametric Distributions Used to Model the Growth Potential at 5 $^{0}$C for Smoked Seafood.

The methodology of application of sensitivity analysis remains exactly the same as the variability under uncertainty realizations case but the dataset used represents uncertainty distributions under various variability conditions. The uncertainty in two of the three uncertain inputs for exposure module of *Listeria monocytogenes* was in the form of a choice among the alternative parametric distributions used to the fit to the data. The variation due to uncertainty in such cases was very small compared to the variation due to variability. Figure 12-7 shows a plot for three parametric probability distribution models fit the data for growth potential at 5 $^{0}$C for smoked seafood. These three distributions are nearly indistinguishable from each other in the central tendency and have only minor differences in the tail relative to the overall range of variability. Thus, the results of a two-dimensional sensitivity analysis in this case are not expected to vary much with regard to different variability iterations. Hence, a two-dimensional analysis aimed at key sources of uncertainty was not conducted for the exposure module. The dose response module has only an uncertainty dimension. Hence only uncertainty analysis was performed on the dose response module.

The inputs are similar among the food categories in the *Listeria monocytogenes* model. The relative importance of inputs is different from one category to other. This is because the parameters of the distributions vary for the same inputs among the food categories. The inputs and outputs of concern were identified in both the exposure and dose response modules. The

Table 12-3. List of Input Variables in Exposure Module

| Name | Unit | Distribution Variability | Distribution Uncertainty | Comments |
|---|---|---|---|---|
| Serving Size | grams | Empirical | None | |
| Initial LM Concentration (at retail) | Log cfu/g | Weighted parametric models (e.g., Beta, Weibull, Triangular, Lognormal) | Selection of alternative parametric model | Choose a parametric model randomly in proportion of the weights |
| Storage Temperature | $^0$C | Empirical | None | |
| Storage Time | day | Beta Pert | Uniform (+-20% for most frequent value and +-50% for maximum value) | 100% correlation between uncertainty distribution for most frequent and maximum value |
| Growth at 5 $^0$C | Log cfu/ day | Weighted parametric model (e.g., Lognormal, Gamma, triangular, Beta) | Selection of alternative parametric model | Choose a parametric model randomly in proportion of the weights |

exposure module has two types of inputs. Some inputs have only variability associated with them, whereas others have both variability and uncertainty. All variables in the dose response module were only uncertain.

Five inputs of interest for each food group were identified in the exposure module. The inputs of interest included serving size in grams, initial LM concentration in log cfu/g, storage temperature in $^0$C, storage time in days, and growth potential at 5 $^0$C in log cfu/day. The output of interest is the dose value corresponding to each meal serving size simulated. The variables along with their distributions are listed in Table 12-3. Serving size has a distribution based on the amounts and frequency of consumption of the food. Initial *Listeria monocytogenes* concentration is at retail and before growth. It has a distribution based on frequency and levels of *Listeria monocytogenes* in ready-to-eat foods. The only other variable in the exposure module is an intermediate variable for maximum log growth that is dependent on the temperature.

Table 12-4 shows the inputs considered in dose response module. The inputs and outputs of interest in dose response considered different sources of uncertainty. Servings per annum and pregnancy rates vary from one food to another but remain constant within a food category. The variables of concern identified are uncertainty in exposure period; uncertainty in virulence which

Table 12-4. List of Input Variables in Dose Response Module

| Name | Unit | Distribution Uncertainty | Source |
|---|---|---|---|
| Dose Adjustment Factor | Log cfu/serving | Empirical | Goal Seek |
| Exposure Period | days | Triangular | Chosen distribution |
| Virulence Susceptibility | Log cfu/serving | Empirical | Virulence simulation |
| Mouse Lethality | Deaths | Empirical | Mouse Experiments |
| Fraction of Population Exposed | NA | Empirical | Exposure module |

considering adjustments for strain and susceptibility virulence; uncertainty in mouse lethality rate because uncertainty in response to a given dose; exposure uncertainty which is due to varying fractions in a dose bin from different uncertainty runs of exposure module; and uncertainty in the adjustment factor as generated by goal seek in each uncertainty iteration. Although the dose adjustment factor is calculated during the simulation, for purpose of gaining insight into whether this parameter is highly sensitive to total risk, it is treated as if it were an input when performing sensitivity analysis. Thus, the adjustment factor is treated as an input just as it would be in case when the dose response model is used for mortality prediction rather than calibration to actual data. The output of interest here is the mortality.

## 12.8 Model Limitations and Modifications

In the process of applying sensitivity analysis to the *Listeria monocytogenes* model several limitations were faced and accordingly modifications were done. This section describes the limitations of the model and the modifications that were done to enable sensitivity analysis.

### 12.8.1 Limitations

A global sensitivity analysis is preferred over local sensitivity so that the analysis can be directly related to the output of decision importance. For food safety risk assessment, the output of decision importance is typically morbidity or mortality. Sensitivity analysis methods typically require a one-to-one mapping of each input to the output. The *Listeria monocytogenes* model is not suitably structured for this type of global sensitivity analysis. Thus only local sensitivity analysis within modules of the model is possible. Understanding code embedded in the Excel spreadsheet is difficult and time consuming. The sequence of operations cannot be easily inferred

by inspection of the sheet. This is a major hindrance in understanding a complex model such as this. The limitations of the *Listeria monocytogenes* model with respect to the application of sensitivity analysis can be characterized in two groups:

- Modularity and Binning
- Coding Limitations

Modularity is a way of organizing a model into sub-divisions. For any sensitivity analysis it is important to distinguish among the parameters that are regarded as inputs, intermediate variable and outputs. The mathematical or stochastic structure of a model is independent of the numerical value of the input assumptions. The inputs and outputs to the model should be in separate modules from the mathematical model. This is not done in *Listeria monocytogenes* model code; the inputs/outputs and growth model are in the same module. In cases where modules have many-to-one mapping between inputs and outputs, the one-to-one correspondence between an input and output is lost. The exposure module of the *Listeria monocytogenes* model allows mapping of several individual meal servings to the same dose bin. After binning, the meal serving that resulted in a particular dose cannot be identified. This causes loss of information. Although this kind of many-to-one mapping is necessary in some cases, the binning can be done while also applying additional methods on an unbinned output in parallel. This will preserve the information as well as allow the binning to be done, although it may be at the cost of additional computational time.

Part of the coding of *Listeria monocytogenes* model was done in MS Excel. This made the understanding of the complex model a tedious job, given the difficulty of inspecting code embedded in a spreadsheet. Therefore, modification of the model and addition of new modules was difficult. For a complex model, implementation of the complete model using a programming language such as Visual Basic would allow for easier understanding, modification and addition of features. The model was recoded by the FDA into Visual Basic during the time of this work, which facilitated application of sensitivity analysis methods.

### 12.8.2 Modifications to Enable Sensitivity Analysis

The Excel-based *Listeria monocytogenes* model was not originally implemented with an objective of supporting sensitivity analysis as extensive as those conducted in this work. Therefore, it was necessary to make modifications to the model in order to facilitate the application of a wide variety of sensitivity analysis methods. During the course of this work,

FDA reimplemented the *Listeria monocytogenes* model using Visual Basic macros. The reimplemented code greatly facilitated the application of sensitivity analysis methods because it was easier to inspect the code, identify model inputs and outputs, and collect data for these variables for use in sensitivity analysis. The latter involved creation of additional worksheets for the purpose of storing data values for inputs and outputs during the course of a probabilistic simulation. In addition, it was necessary to modify the code in order to extract values for intermediate variables of interest. For example, in the dose response module goals seek was performed to get the corrected dose and mortality. The inputs used to get these doses were not stored. Thus to get the inputs, code was inserted in the model.

To apply mathematical sensitivity analysis methods, a module for each of NRSA and differential sensitivity analysis was inserted in the exposure and dose response parts. The insertion of the NRSA and differential sensitivity analysis module did not change the underlying structure of the model. For conditional sensitivity analysis, uncertain input variables were fixed at point estimate values. This was achieved in two ways: (a) when the uncertainty was in form of a distribution, a probability value of 0.5 was used to get the median as the desired point estimate; (b) when uncertainty was in terms of selection of a distribution, then a weighted average of all distribution was taken to give the point estimate.

In the original *Listeria monocytogenes* model, the dose response module stored the number of deaths at each uncertain step but did not store the corresponding input values. Since mortality calculations are done inside the Excel spread sheet, code was added to create a dataset. For each uncertain run, an adjustment factor was generated by running "goal seek" to match the predicted and actual number of deaths from all food categories in a given sub-population. The dataset contains the inputs and mortality predicted for each food category after dose adjustment using the adjustment factor.

## 12.9 Generation of Datasets for Sensitivity Analysis

The sensitivity analysis methods use datasets specific to food categories and based on the input assumptions for variability and uncertainty. For example, mathematical methods need point estimate values for each input. The statistical methods need the input and output values with one-to-one correspondence for each iteration in a Monte Carlo simulation. This section discusses the assumptions and number of iterations used in generation of data.

Table 12-5. Minimum, Median and Maximum Values of Input variables for the Five Food Categories

| Input[(a)] | Point Estimates | Pâtés | Smoked Sea Food | Fresh Soft Cheese | Milk | Deli Salad |
|---|---|---|---|---|---|---|
| Serving Size in grams | Min | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | Median | 57.0 | 57.0 | 62.0 | 244 | 115 |
| | Max | 454 | 142 | 246 | 3900 | 1410 |
| Initial LM Concentration in log cfu/g | Min | -3.0 | -2.5 | -2.7 | -3.9 | -3.4 |
| | Median | -0.4 | 0.2 | -1.0 | -1.7 | -1.6 |
| | Max | 9.0 | 9.0 | 9.0 | 8.9 | 5.2 |
| Storage temperature in $^0$C | Min | 0.0 | 0.0 | 0.0 | 0.0 | 0.0 |
| | Median | 4.5 | 4.5 | 4.5 | 4.5 | 4.5 |
| | Max | 21 | 21 | 21 | 21 | 21 |
| Storage time in days | Min | 0.8 | 0.6 | 0.5 | 0.6 | 0.6 |
| | Median | 8.4 | 5.0 | 4.0 | 4.2 | 3.7 |
| | Max | 45 | 30 | 30 | 17 | 14 |
| Growth at 5 $^0$C in log cfu/day | Min | -0.3 | 0.0 | -0.4 | 0.0 | -0.5 |
| | Median | 0.3 | 0.1 | 0.1 | 0.3 | -0.1 |
| | Max | 0.7 | 0.4 | 0.5 | 0.8 | 0.2 |

(a) For each input the minimum, median and maximum values are shown.

To get a one-to-one correspondence between inputs and the output, the output was recorded before binning in the exposure module. Separate simulations were made for variability only and both variability and uncertainty cases. To enable application of ANOVA, discrete levels were created from the dataset generated from simulation. To calculate the minimum, median and maximum values of each input in the exposure module a simulation with 3000 variability iterations and 250 uncertainty iterations were done for each food category.

The presence of *Listeria monocytogenes* cannot be detected in food below a detection limit. This detection limit was set as a truncation value. For the cases where the dose value was below the truncation value, the input data corresponding to that dose were discarded. Thus all doses used in sensitivity analysis had values above the truncation value and the dataset size was much smaller than the original 3000 variability and 300 uncertain iteration dataset. This process of selection of data helped in concentrating the analysis on the tail or high end of exposure values where Listeriosis is likely to occur. The minimum, median and maximum values among all the values generated were calculated for each of the five food categories. These are shown in Table 12-5. For example, the minimum, median and maximum values of serving size for smoked seafood are 0.0 g, 57.0 g, and 142 g, respectively. The 0.0 g of serving size corresponds to the

lower end of the lognormal distribution. Based on the median value for each input in the Table 12-5, the lower and upper differential values for inputs were calculated for use in DSA. The lower value was given by 99 percent of the median and the upper value by 101% of the median value. Thus differential sensitivity was tested under a perturbation of 1% on either side of the median.

The output of the exposure module in the original model is the fraction of doses in each of the half log dose bins. The aggregation of the data to bins causes loss of information regarding the input conditions that resulted in a particular dose. Thus the output for analysis was chosen as the direct dose value, calculated after growth at post-retail storage. The cases considering only variability used 30,000 iterations. For cases where both variability and uncertainty were considered, the number of variability iterations was 3000 and the number of uncertainty iterations was 250. Large portions of these did not appear in actual analysis because of truncation. Typically ten percent of the samples were above the truncation value.

To simulate a randomly selected person from a population a simulation was run with 1000 variability iterations and 200 uncertainty iterations and the dataset for different uncertainty realizations were combined into one dimension.

In order to perform ANOVA, each factor should be divided into discrete levels. For continuous inputs, levels were defined by splitting the domain of values into ranges based upon the cumulative distribution function of the input. In particular, levels were defined based upon the lower tail, middle region, and upper tail of the distribution of each input. For each input, Level 1 was defined to include the values of the input up to the $20^{th}$ percentile of the distribution. Level 2 was defined to include values between the $20^{th}$ percentile and $80^{th}$ percentile. Level 3 was defined to include values above the $80^{th}$ percentile.

For the dose response module, the minimum, median and maximum are calculated in manner similar to exposure part. The number of simulations considered was 4000 uncertainty realizations. For virulence susceptibility, mouse lethality and fraction of population exposed, the random number associated with uncertainty was directly chosen as the values of the corresponding inputs in the input/output dataset. This was done because of the specific methodology of matrix adjustment for virulence and host susceptibility left no reasonable way to extract the value of variable itself under uncertainty (CFSAN, 2001). The values of minimum, median and maximum values for the five food categories of concern are shown in Table 12-6.

Table 12-6. Minimum, Median and Maximum Values of Dose Response Module Input Variables for the Five Food Categories of Deli Salad, Fresh Soft Cheese, Milk, Pâtés and Smoked Seafood.

| Name | Minimum | Median | Maximum |
|---|---|---|---|
| Dose Adjustment Factor | 5.11 | 7.35 | 11.3 |
| Exposure Period | 1.32 | 11.9 | 29.8 |
| Virulence Susceptibility | 0.0 | 0.48 | 1.00 |
| Mouse Lethality | 0.0 | 0.48 | 1.00 |
| Fraction of Population Exposed | 0.0 | 0.48 | 1.00 |

For example, the minimum, median and maximum values of dose adjustment factor are 5.11, 7.35, and 11.3, respectively. The dose adjustment factor is estimated based upon the cumulative impact of all food groups. Thus, the distribution shown is inferred from model results and was not specified exogenously.

This chapter summarized the draft of *Listeria monocytogenes* risk assessment. Model limitations, modifications and data generation were also described. Chapters 13 and 14 document the application of NRSA and differential sensitivity analysis methods, respectively. Chapters 15, 16 and 17 document the application of the regression analysis, ANOVA and CART, respectively. Chapter 18 presents scatter plots and Chapter 19 summarizes and compares the various methods applied to the *Listeria monocytogenes* model.

# 13 APPLICATION OF NOMINAL RANGE SENSITIVITY ANALYSIS TO THE *LISTERIA MONOCYTOGENES* MODEL

The purpose of this chapter is to apply NRSA to the *Listeria monocytogenes* model. NRSA is discussed in Section 2.1.1. The *Listeria monocytogenes* model is discussed in Chapter 13. NRSA is most applicable to linear models. The *Listeria monocytogenes* model is nonlinear. Therefore, a key consideration in this chapter is to evaluate whether the linear basis for NRSA is sufficiently robust that it can generate useful insights when applied to a nonlinear model. This chapter is divided into three major sections. The first section focuses on the application of NRSA to the exposure module of the *Listeria monocytogenes* model. The second section focuses on the application of NRSA to the dose-response module of the *Listeria monocytogenes* model. The third section summarizes the main findings from the application of NRSA to the *Listeria monocytogenes* model.

## 13.1 Application of NRSA to the Exposure Module

NRSA was applied to the exposure module for five selected food groups, including deli salad, fresh soft cheese, milk, pâtés, and smoked seafood. For each food group, the inputs of interest include serving size in grams, initial LM concentration in log cfu/g, storage temperature in $^0$C, storage time in days, and growth potential at 5 $^0$C in log cfu/day. The output of interest is the dose value corresponding to each meal serving size simulated. In order to perform NRSA, a nominal point estimate and a range of values must be identified for each input. The median, minimum, and maximum values for all inputs for all food groups are given in Table12-5. The sensitivity indices of the five inputs for each food group were calculated, and the results are summarized in Table 13-1.

The most important input for each of the five food categories was found to be the initial LM concentration. The serving size and the storage temperature are the second and third most important inputs for all food categories. Storage time and growth at 5 $^0$C are the least two important inputs for all foods. There is no measure of statistical significance for the sensitivity index of NRSA. However, qualitative insight into the robustness of the rankings can be obtained by comparing the magnitude of the sensitivity index for closely ranked inputs. For example, the sensitivity indices for initial LM concentration and serving size are 95 and 71, respectively, for deli salad and are 6.0 and 5.2, respectively, for smoked seafood. These results imply that the rank order of these two inputs is probably more robust in the case of deli salad than for smoked

seafood, because in the latter case the sensitivity indices are closer in numerical values than in the former case. For all food groups, the second highest ranked variable has a sensitivity index substantially larger than the third ranked variable. For example, for deli salad the magnitude of the sensitivity index of the second highest ranked input of 71 is more than a factor of two larger than that of the third ranked input, which had a magnitude of 32. Similarly, for fresh soft cheese, the sensitivity index of the second ranked input is nearly a factor of four larger than that for the third ranked input.

The sensitivity indices of the two lowest ranked inputs typically differ from each other by a factor of 1.3 to 2.1 among the five foods. However, the sensitivity index of the fourth ranked input differs from that of the top ranked input by a factor of 2.7 to 7.0. Thus, the fourth ranked input is substantially less important than the top ranked input, and the fourth and fifth ranked inputs are approximately comparable in importance for at least some cases, such as for milk.

NRSA is based upon a linearity assumption and does not take into account simultaneous variation in all inputs. Thus, a key question regarding this method is whether the rankings are comparable to those of methods, such as ANOVA, that account for simultaneous variation in multiple inputs and do not assume a linear functional relationship. The results of the application of NRSA and ANOVA to the *Listeria monocytogenes* model are given in Table 14-1 and Table 16-5. Both NRSA and ANOVA identify initial LM concentration as the first ranked input. However, for all food categories the inputs second ranked by the two methods differ. ANOVA ranks growth at 5 $^0$C as second for deli salad and fresh soft cheese, whereas NRSA ranks the growth at 5 $^0$C as for deli salad and fresh soft cheese. Also, for milk, pâtés and smoked seafood, NRSA ranks storage temperature as third compared to the second rank given by ANOVA. The difference in ranking given by NRSA and methods such as ANOVA can be attributed to the non-linear model response to growth at 5 $^0$C and storage temperature and the considerable variation in responses at different point estimates. This is evident from the scatter plots shown in Chapter 18.

Table 13-1. Results of Application of Nominal Range Sensitivity Analysis to the *Listeria monocytogenes* Exposure Module for Deli Salad, Fresh Soft Cheese, Milk, Pâtés, and Smoked Seafood.

| Variable Name | Nominal Range Sensitivity Index | Rank Within the Food Category |
|---|---|---|
| **Deli Salad** | | |
| 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 |
| **Fresh Soft Cheese** | | |
| Serving Size (g) | 11.9 | 2 |
| Initial LM Concentration (log cfu/g) | 15.4 | 1 |
| Storage temperature ($^0$C) | 3.1 | 3 |
| Storage time (days) | 1.4 | 5 |
| Growth at 5 $^0$C (log cfu/day) | 3.0 | 4 |
| **Milk** | | |
| Serving Size (g) | 8.0 | 2 |
| Initial LM Concentration (log cfu/g) | 10.6 | 1 |
| Storage temperature ($^0$C) | 4.5 | 3 |
| Storage time (days) | 2.1 | 4 |
| Growth at 5 $^0$C (log cfu/day) | 1.6 | 5 |
| **Pâtés** | | |
| Serving Size (g) | 4.5 | 2 |
| Initial LM Concentration (log cfu/g) | 4.8 | 1 |
| Storage temperature ($^0$C) | 2.1 | 3 |
| Storage time (days) | 1.1 | 5 |
| Growth at 5 $^0$C (log cfu/day) | 1.8 | 4 |
| **Smoked Seafood** | | |
| Serving Size (g) | 5.2 | 2 |
| Initial LM Concentration (log cfu/g) | 6.0 | 1 |
| Storage temperature ($^0$C) | 2.4 | 3 |
| Storage time (days) | 1.2 | 4 |
| Growth at 5 $^0$C (log cfu/day) | 0.7 | 5 |

## 13.2 Application of NRSA to the Dose Response Module

NRSA was applied to the dose response module for five selected food groups, including deli salad, fresh soft cheese, milk, pâtés, and smoked seafood. For each food group, the inputs of interest include uncertainty in dose adjustment factor, exposure period, virulence susceptibility uncertainty, and mouse lethality uncertainty and population exposure. The output of interest is mortality ascertained to the food group in the neonatal sub-population. In order to perform NRSA, a nominal point estimate and a range of values were identified for each input. The median, minimum, and maximum values for all inputs for all food groups are given in Table 12-6. As seen in the table the minimum, median and maximum values for inputs are same across all food categories. Thus, same result is obtained for all food categories and is summarized in Table 14-2.

For all of the five food categories, the top input was mouse lethality. The sensitivity index for mouse lethality and fraction of population exposed are very close, thus these two inputs are likely to be of approximately comparable importance. Virulence susceptibility, exposure period, and dose adjustment factor are ranked as the third, fourth, and fifth, respectively, important inputs. However, the sensitivity indices for the fourth ranked input is small compared to third ranked input. Thus, although there is some ambiguity regarding the most important input, there is no ambiguity regarding which input is of the third most importance. Moreover, the sensitivity index of the fifth ranked input is orders of magnitude smaller than for the fourth ranked input. In summary, for all of the five food categories, the mortality is most sensitive to the mouse lethality and fraction of population exposed, moderately sensitive to virulence susceptibility, and is comparatively insensitive to the exposure period and the dose adjustment factor.

Table 13-2. Results of Application of Nominal Range Sensitivity Analysis to the *Listeria monocytogenes* Dose Response Module for Deli Salad, Fresh Soft Cheese, Milk, Pâtés, and Smoked Seafood

| Variable Name | Nominal Range Sensitivity Index | Rank Within the Food Category |
|---|---|---|
| Dose Adjustment Factor | 4.47E-07 | 5 |
| Exposure Period | 4.2E-03 | 4 |
| Virulence Susceptibility | 55.5E-03 | 3 |
| Mouse Lethality | 72.8E-03 | 1 |
| Fraction of Population Exposed | 70.2E-03 | 2 |

## 13.3 Summary of the Result of Application of NRSA

NRSA consistently ranked the top three variables for exposure as initial LM concentration, serving size and storage temperature and, the top three variables for dose response as uncertainty in mouse lethality, virulence susceptibility and fraction of population exposed. The sensitivity indices for the top two variables in exposure module were substantially different from the others. The sensitivity indices for the top three variables in dose response were substantially different from the others. Thus NRSA consistently separated out the sensitive variables from less important ones. The performance of NRSA is compared with other sensitivity analysis methods in Chapter 19.

# 14 APPLICATION OF DIFFERENTIAL SENSITIVITY ANALYSIS TO THE *LISTERIA MONOCYTOGENES* MODEL

The purpose of this chapter is to apply Differential Sensitivity Analysis (DSA) to the *Listeria monocytogenes* model. DSA is discussed in Section 2.1.2. The *Listeria monocytogenes* model is discussed in Chapter 12. DSA is a local sensitivity analysis method that evaluates the effect of small perturbation in each input independently with respect to a single point. The reference point for analysis is based upon the median values of all inputs. DSA does not consider the full range of possible variation in each input. A key consideration in this chapter is to evaluate whether sensitivity indices with respect to the median adequately represents the sensitivity over a full range of variation of the variables in a non-linear model.

This chapter is divided into three major sections. The first section focuses on the application of DSA to the exposure module of the *Listeria monocytogenes* model. The second section focuses on the application of DSA to the dose-response module of the *Listeria monocytogenes* model. The third section summarizes the main findings from the application of DSA to the *Listeria monocytogenes* model.

## 14.1 Application of Differential Sensitivity Analysis to the Exposure Module

DSA was applied to the exposure module for five selected food groups, including deli salad, fresh soft cheese, milk, pâtés, and smoked seafood. For each food group, the inputs of interest include serving size in grams, initial LM concentration in log cfu/g, storage temperature in $^0$C, storage time in days, and growth potential at 5 $^0$C in log cfu/day. The output of interest is the dose value corresponding to each meal serving size simulated. In order to perform DSA, a nominal point estimate and a delta range of values around the nominal must be identified for each input.

The median values are listed in Table 12-5. The median, 99% of median and, 101% of median values were considered for analysis. The sensitivity indices of the five inputs for each food group were calculated, and the results are summarized in Table 14-1.

The initial LM concentration is the most sensitive input and has a sensitivity index substantially larger, by a factor of 20 or more, than that of the second ranked input for each food category. Storage temperature has the second highest sensitivity index. However, the sensitivity indices for the second through fifth ranked inputs typically do not differ substantially from each

Table 14-1. Results of Application of Differential Sensitivity Analysis to the *Listeria monocytogenes* Exposure Module for Deli Salad, Fresh Soft Cheese, Milk, Pâtés, and Smoked Seafood

| Variable Name | Differential Sensitivity Index | Rank Within the Food Category |
|---|---|---|
| **Deli Salad** | | |
| Serving Size (g) | 2.7 | 3 |
| Initial LM Concentration (log cfu/g) | 730 | 1 |
| Storage temperature ($^0$C) | 3.2 | 2 |
| Storage time (days) | 2.0 | 4 |
| Growth at 5 $^0$C (log cfu/day) | 2.0 | 4 |
| **Fresh Soft Cheese** | | |
| Serving Size (g) | 0.4 | 2 |
| Initial LM Concentration (log cfu/g) | 63 | 1 |
| Storage temperature ($^0$C) | 0.3 | 3 |
| Storage time (days) | 0.2 | 4 |
| Growth at 5 $^0$C (log cfu/day) | 0.2 | 4 |
| **Milk** | | |
| Serving Size (g) | 0.3 | 5 |
| Initial LM Concentration (log cfu/g) | 150 | 1 |
| Storage temperature ($^0$C) | 0.9 | 2 |
| Storage time (days) | 0.6 | 3 |
| Growth at 5 $^0$C (log cfu/day) | 0.6 | 3 |
| **Pâtés** | | |
| Serving Size (g) | 0.1 | 5 |
| Initial LM Concentration (log cfu/g) | 18 | 1 |
| Storage temperature ($^0$C) | 0.9 | 2 |
| Storage time (days) | 0.6 | 3 |
| Growth at 5 $^0$C (log cfu/day) | 0.6 | 3 |
| **Smoked Seafood** | | |
| Serving Size (g) | 0.2 | 3 |
| Initial LM Concentration (log cfu/g) | 23 | 1 |
| Storage temperature ($^0$C) | 0.3 | 2 |
| Storage time (days) | 0.2 | 3 |
| Growth at 5 $^0$C (log cfu/day) | 0.2 | 3 |

other for any food category. These results imply that the dose is sensitive to small perturbation in initial LM concentration but not to small perturbations in the other inputs.

The results from DSA differ from that of other methods, such as NRSA, because DSA does not consider the full range of variation in the inputs. For example, although both methods lead to the identification of the initial LM concentration as the most sensitive input, NRSA also implied that serving size and storage temperature had substantial sensitivity. In contrast, with DSA, there was no substantial distinction between serving size or storage temperature compared to storage time and growth at 5 $^{o}$C. All four of these inputs had very small sensitivity indices compared to the top-ranked input. Thus, as expected, DSA does not provide the same insights regarding sensitivity as does a method that considers the full range of variation, and not just small perturbations, for each input.

## 14.2 Application of Differential Sensitivity Analysis to the Dose Response Module

DSA was applied to the dose response module for five selected food groups, including deli salad, fresh soft cheese, milk, pâtés, and smoked seafood. For each food group, the inputs of interest include uncertainty in dose adjustment factor (log cfu), exposure period, virulence susceptibility uncertainty, mouse lethality uncertainty, and population exposure (log cfu). The output of interest is mortality ascertained to the food group in the sub-population. In order to perform DSA, a nominal point estimate and a delta range of values around the nominal must be identified for each input. The median values are listed in Table 13-5. The median values for the inputs were same across all food categories. The median, 99% of median and, 101% of median values for all inputs were considered for analysis. The sensitivity indices of the five inputs are summarized in Table 14-2.

The rankings assigned to the inputs are the same across all food categories. The fraction of population exposed is the most sensitive input, followed in order by mouse lethality, virulence susceptibility, exposure period, and dose adjustment factor. The ratio of the sensitivity indices for the highest versus lowest ranked inputs varies by a factor of 7.4 to 18.5 among the food categories. The indices among the five inputs appear to be sufficiently different from each other such that there is little ambiguity regarding the ranks assigned. For example, the sensitivity index of the top ranked input is a factor of 1.53 greater than that for the second ranked input.

Table 14-2. Results of Application of Differential Sensitivity Analysis to the *Listeria monocytogenes* Dose Response Module for Deli Salad, Fresh Soft Cheese, Milk, Pâtés, and Smoked Seafood

| Variable Name | Differential Sensitivity Index | Rank Within the Food Category |
|---|---|---|
| Dose Adjustment Factor | 0.99 | 5 |
| Exposure Period | 1.46 | 4 |
| Virulence Susceptibility | 4.50 | 3 |
| Mouse Lethality | 5.68 | 2 |
| Fraction of Population Exposed | 8.68 | 1 |

Compared to NRSA, DSA provides a similar although not identical ranking of the inputs. The rank order of the top two inputs is reversed. Because DSA considers the effects of only small perturbations in the inputs, and not the effect of variation over the range of likely values, the differences in results of the two methods are attributed to differences in the ranges included for each input.

## 14.3 Summary of the Result of Application of Differential Sensitivity Analysis

Differential sensitivity analysis consistently ranked initial LM concentration as the most sensitive input variable in the exposure module. The other inputs are substantially less sensitive by comparison. The indices for dose response were well separated for the inputs within each of food categories. The relative ranking of inputs was the same for different food categories. The uncertainty in the fraction of population exposed was the most sensitive dose response input.

DSA is a useful method for understanding the model response to a small perturbation in each of the inputs. However, because each input has a different relative range of variation, and because this information is not included in DSA, the results from DSA may differ from those obtained by methods in which the range of variation is accounted for. Thus, in presenting results with DSA, it is appropriate to make inferences regarding the sensitivity of the model to each input. However, it is difficult to make inferences regarding the importance of each input unless the likely range of variation of the input is also accounted for. Thus, DSA could be a useful tool, for example, in understanding the local behavior of a model but may not be useful, especially for nonlinear models, in making inferences regarding which inputs are the most important ones.

A comparison of results with DSA and other methods, such as NRSA, can provide insight into whether it is the structure of the model or the range of variation in the input that is contributing the most to the results of NRSA. For example, in the exposure module, the results

of DSA show that the model is far more sensitive to small changes in the initial LM concentration than for any other input. Thus, to the extent that the model response to initial LM concentration is linear or monotonic, it would not take a very large range of variation in LM concentration to contribute to a large range of model response. The results of NRSA imply that the initial LM concentration is the most important input, and that serving size is the second most important input. By comparison, the results of DSA imply that the model response to a small perturbation in serving size is much smaller than for several other inputs. Therefore, it must be the case that there is a large range of variation in serving size in order for NRSA to indicate that serving size is more important than all other inputs aside from the initial LM concentration. Thus, in this case, it is the range of variation in serving size, as opposed to the sensitivity of the model to a small perturbation in serving size that contributes to the importance of this particular input.

# 15 APPLICATION OF REGRESSION ANALYSIS TO THE *LISTERIA MONOCYTOGENES* MODEL

The purpose of this chapter is to apply regression analysis to the *Listeria monocytogenes* model. Regression analysis is discussed in Section 2.2.2. The *Listeria monocytogenes* model is discussed in Chapter 12. Regression analysis assumes a model and involves fitting the model to the data generated for inputs and outputs during Monte Carlo simulation. The *Listeria monocytogenes* model is nonlinear. Although it is possible to use non-linear basis functions in regression analysis, regression analysis is typically done using linear models. Therefore, a key consideration in this chapter is to evaluate whether the non-linearity in *Listeria monocytogenes* model is adequately captured by linear regression analysis.

In order to apply regression analysis to different modules of the *Listeria monocytogenes* food safety risk assessment model, SAS© Version 8.02 was used. This software package has the ability to perform regression analysis by using available procedures on a dataset. The macro feature in SAS allows programs to be written to automate the process of application of available procedures on multiple dataset.

This chapter is divided into three major sections. The first section focuses on the application of regression analysis to the exposure module of the *Listeria monocytogenes* model. The second section focuses on the application of regression analysis to the dose-response module of the *Listeria monocytogenes* model. The third section summarizes the main findings from the application of regression analysis to the *Listeria monocytogenes* model.

## 15.1 Application of Regression Analysis to the Exposure Module

Regression analysis was applied to the exposure module for five selected food groups, including deli salad, fresh soft cheese, milk, pâtés, and smoked seafood. For each food group, the inputs of interest include serving size in grams, initial LM concentration in log cfu/g, storage temperature in $^0$C, storage time in days, and growth potential at 5 $^0$C in log cfu/day. The output of interest is the dose value corresponding to each meal serving size simulated.

Regression analysis was applied on three types of datasets. A dataset considering only variability was analyzed. A mixed dataset where both variability and uncertainty are present but not distinguished was analyzed to consider a randomly selected individual. Finally, variability in dataset was analyzed under a number of uncertainty realizations using a two-dimensional simulation. The details of dataset generation are described in Section 12.9.

The following subsection focuses on application of regression analysis based upon only variability in inputs and the corresponding estimated variability in dose. Section 15.1.2 focuses upon application of regression analysis when both variability and uncertainty are considered together. The purpose of conducting both analyses is to gain insight into how the contribution of uncertainty changes the results of sensitivity analysis compared to a case in which only variability is considered. Section 15.1.3 gives the analysis where variability is analyzed under various uncertain realizations. This case enables evaluation of the confidence in rankings under different uncertainty realization. The parameter estimates, 95 percent parameter confidence intervals and relative rankings are specified in the results of each analysis.

### 15.1.1 Sensitivity Analysis of Exposure Module Based Upon Variability Only

This section covers application of regression analysis to the exposure module when only variability was considered during simulation. Since the original exposure module considers both variability and uncertainty, the exposure module was modified to remove the uncertainty dimension. The uncertainty dimension was fixed to median or most likely values as described in Section 12.11. Regression analysis was performed for each of five selected food categories. The results of regression analysis for each input are shown in Table 15-1.

For each food category, the five exposure models inputs are assigned a sensitivity rank based upon the magnitude of their respective parameter estimates. For example, for deli salad, the initial LM concentration has a statistically significant parameter estimate of 0.81, which is a factor of 2.3 greater than the input with the second highest parameter estimate. Because the confidence interval for the parameter estimates does not overlap with that of the second ranked input, the initial LM concentration is significantly more sensitive then the second ranked input of growth rate at 5 $^{0}$C. Thus, initial LM concentration is inferred to be the most sensitive input for this food category.

The parameter estimate of the second ranked input is a factor of 1.3 greater than that for the third ranked input and the two are significantly different. Thus, the first, second and third ranked inputs have parameter estimates that differ substantially. Therefore, these three inputs are of descending importance. The smallest parameter estimate of 0.10 for storage time is statistically significant but is substantially smaller than that of the fourth ranked input. The implication is that storage time by itself contributes little to variation in dose.

Table 15-1. Parameter Estimates and Within Food Category Rankings for Variability Only for the *Listeria monocytogenes* Exposure Model Inputs for Deli Salad, Fresh Soft Cheese, Milk, Pâtés, and Smoked Seafood

| Variable Name | Parameter Estimate | 95% Confidence Interval | Rank Within The Food Category | Statistically Significant |
|---|---|---|---|---|
| **Deli Salad ($R^2$ = 0.80)** | | | | |
| Serving Size (g) | 0.27 | 0.26, 0.29 | 3 | Yes |
| Initial LM Concentration (log cfu/g) | 0.81 | 0.79, 0.82 | 1 | Yes |
| Storage temperature ($^0$C) | -0.26 | -0.27, -0.24 | 4 | Yes |
| Storage time (days) | -0.10 | -0.11, -0.09 | 5 | Yes |
| Growth at 5 $^0$C (log cfu/day) | 0.35 | 0.34, 0.37 | 2 | Yes |
| **Fresh Soft Cheese ($R^2$ = 0.87)** | | | | |
| Serving Size (g) | 0.25 | 0.24, 0.26 | 3 | Yes |
| Initial LM Concentration (log cfu/g) | 0.88 | 0.87, 0.89 | 1 | Yes |
| Storage temperature ($^0$C) | 0.13 | 0.12, 0.14 | 4 | Yes |
| Storage time (days) | 0.07 | 0.06, 0.08 | 5 | Yes |
| Growth at 5 $^0$C (log cfu/day) | 0.36 | 0.35, 0.38 | 2 | Yes |
| **Milk ($R^2$ = 0.86)** | | | | |
| Serving Size (g) | 0.17 | 0.15, 0.19 | 5 | Yes |
| Initial LM Concentration (log cfu/g) | 0.74 | 0.72, 0.75 | 1 | Yes |
| Storage temperature ($^0$C) | 0.57 | 0.56, 0.59 | 2 | Yes |
| Storage time (days) | 0.24 | 0.23, 0.26 | 4 | Yes |
| Growth at 5 $^0$C (log cfu/day) | 0.27 | 0.25, 0.28 | 3 | Yes |

(Continued on next page.)

Table 15.1. Continued

| Variable Name | Parameter Estimate | 95% Confidence Interval | Rank Within The Food Category | Statistically Significant |
|---|---|---|---|---|
| **Pâtés ($R^2$ = 0.84)** | | | | |
| Serving Size (g) | 0.10 | 0.09, 0.12 | 5 | Yes |
| Initial LM Concentration (log cfu/g) | 0.64 | 0.62, 0.65 | 2 | Yes |
| Storage temperature ($^0$C) | 0.69 | 0.68, 0.71 | 1 | Yes |
| Storage time (days) | 0.29 | 0.27, 0.30 | 3 | Yes |
| Growth at 5 $^0$C (log cfu/day) | 0.18 | 0.16, 0.19 | 4 | Yes |
| **Smoked Seafood ($R^2$ = 0.92)** | | | | |
| Serving Size (g) | 0.16 | 0.15, 0.16 | 4 | Yes |
| Initial LM Concentration (log cfu/g) | 0.90 | 0.89, 0.91 | 1 | Yes |
| Storage temperature ($^0$C) | 0.35 | 0.34, 0.36 | 2 | Yes |
| Storage time (days) | 0.17 | 0.17, 0.18 | 3 | Yes |
| Growth at 5 $^0$C (log cfu/day) | 0.12 | 0.12, 0.13 | 5 | Yes |

The 95 percent confidence intervals for all parameter estimates are non-overlapping except for serving size and storage temperature where the magnitudes overlap. Hence the ranking of the third and fourth inputs are not robust. Although the storage temperature and serving size are of approximately same importance, a decrease in serving size will decrease the dose whereas a decrease in storage temperature will increase the dose. Thus, the coefficients of these two inputs have different signs. This observation is specific only for deli salad and is biologically plausible. In particular deli salads may contain acidic components that suppress microbial growth more effectively at higher temperature than at lower temperature (CFSAN, 2001).

In comparing results for different food categories, it is apparent that the initial LM concentration is the top ranked input in four of the five cases. In these cases, the parameter estimates for the initial LM concentration are larger by a factor of 1.3 to 2.5 than for the second ranked inputs. For pâtés, where storage temperature was ranked first, the coefficient for the initial LM concentration was a factor of 1.1 smaller and ranked second. The second most important input varies from one food category to another. For milk and smoked seafood, storage

temperature was ranked second with parameter estimates a ratio of 1.3 and 2.6 smaller than the top ranked input. For deli salad and fresh soft cheese, growth at 5 $^0$C was ranked second with parameter estimates a factor of 2.3 and 2.4 smaller than the top ranked input. Storage time and serving size are ranked for two foods.

The parameter estimates for all of the inputs are statistically significant for all food groups. In general, the confidence intervals for the regression coefficients are sufficiently narrow that there is no overlap between confidence intervals for the five inputs for a given food group with only two exceptions. As previously noted for deli salad, the magnitude of the confidence intervals overlap for serving size and storage temperature. For milk, the confidence intervals overlap for storage time and growth at 5 $^0$C. In both of these cases, there is ambiguity regarding which of the inputs should be ranked third versus fourth. However, with these exceptions, the rankings of all other inputs are considered to be robust.

The form of regression analysis used in this study is normalized linear regression analysis, which assumes a linear model. The high $R^2$ values indicate that the response is approximately linear. The $R^2$ values for the regression models were above 0.80 for each of the five food categories. Thus the linear model gave a good fit. However, the output and some of the inputs were in log scale indicating that the relationships were linear with respect to the log scales used in some cases.

### 15.1.2 Sensitivity Analysis of the Exposure Module Based Upon Co-mingled Variability and Uncertainty

This section covers application of regression analysis to the exposure module when both variability and uncertainty are present in the model but are not distinguished. Regression analysis was performed for each of five selected food categories. The details of dataset generation are described in Section 12.12. The results of regression analysis for each input are shown in Table 15-2.

For each food category, the five exposure models inputs are assigned a sensitivity rank based upon the magnitude of their respective parameter estimates. For example, for deli salad, the initial LM concentration has a parameter estimate of 0.80, which is a factor of 2.1 greater than the input with the second highest parameter estimate. Furthermore, the parameter estimate for the initial LM concentration is statistically significant. Thus, initial LM concentration is inferred to be the most sensitive input for this food category, and it is substantially more

Table 15-2. Parameter Estimates and Within Food Category Rankings for Co-Mingled Variability and Uncertainty for the *Listeria monocytogenes* Exposure Model Inputs for Deli Salad, Fresh Soft Cheese, Milk, Pâtés, and Smoked Seafood

| Variable Name | Parameter Estimate | 95% Confidence Interval | Rank Within The Food Category | Statistically Significant |
|---|---|---|---|---|
| **Deli Salad ($R^2$ = 0.82)** | | | | |
| Serving Size (g) | 0.27 | 0.26, 0.27 | 3 | Yes |
| Initial LM Concentration (log cfu/g) | 0.80 | 0.79, 0.80 | 1 | Yes |
| Storage temperature ($^0$C) | -0.35 | -0.36, -0.35 | 4 | Yes |
| Storage time (days) | -0.12 | -0.13, -0.12 | 5 | Yes |
| Growth at 5 $^0$C (log cfu/day) | 0.39 | 0.38, 0.39 | 2 | Yes |
| **Fresh Soft Cheese ($R^2$ = 0.90)** | | | | |
| Serving Size (g) | 0.21 | 0.21, 0.22 | 3 | Yes |
| Initial LM Concentration (log cfu/g) | 0.92 | 0.91, 0.92 | 1 | Yes |
| Storage temperature ($^0$C) | 0.11 | 0.10, 0.11 | 5 | Yes |
| Storage time (days) | 0.13 | 0.12, 0.13 | 4 | Yes |
| Growth at 5 $^0$C (log cfu/day) | 0.31 | 0.30, 0.31 | 2 | Yes |
| **Milk ($R^2$ = 0.88)** | | | | |
| Serving Size (g) | 0.16 | 0.15, 0.16 | 5 | Yes |
| Initial LM Concentration (log cfu/g) | 0.77 | 0.76, 0.77 | 1 | Yes |
| Storage temperature ($^0$C) | 0.46 | 0.46, 0.47 | 2 | Yes |
| Storage time (days) | 0.25 | 0.24, 0.25 | 3 | Yes |
| Growth at 5 $^0$C (log cfu/day) | 0.21 | 0.20, 0.21 | 4 | Yes |

(continued on next page.)

Table 15-2. Continued

| Variable Name | Parameter Estimate | 95% Confidence Interval | Rank Within The Food Category | Statistically Significant |
|---|---|---|---|---|
| **Pâtés ($R^2 = 0.81$)** | | | | |
| Serving Size (g) | 0.06 | 0.06, 0.07 | 5 | Yes |
| Initial LM Concentration (log cfu/g) | 0.63 | 0.63, 0.64 | 1 | Yes |
| Storage temperature ($^0$C) | 0.56 | 0.56, 0.57 | 2 | Yes |
| Storage time (days) | 0.23 | 0.23, 0.24 | 4 | Yes |
| Growth at 5 $^0$C (log cfu/day) | 0.24 | 0.23, 0.24 | 3 | Yes |
| **Smoked Seafood ($R^2 = 0.83$)** | | | | |
| Serving Size (g) | 0.09 | 0.09, 0.10 | 5 | Yes |
| Initial LM Concentration (log cfu/g) | 0.87 | 0.86, 0.87 | 1 | Yes |
| Storage temperature ($^0$C) | 0.30 | 0.29, 0.30 | 2 | Yes |
| Storage time (days) | 0.10 | 0.09, 0.10 | 4 | Yes |
| Growth at 5 $^0$C (log cfu/day) | 0.12 | 0.11, 0.12 | 3 | Yes |

sensitive then the second ranked input of growth rate at 5 $^0$C. The parameter estimate of the second ranked input is a factor of 1.1 greater than that for the third ranked input and 1.4 greater than that for fourth ranked input. Thus, the second, third and fourth ranked inputs have parameter estimates that do not differ substantially. Therefore, these three inputs may be of approximately comparable importance. The smallest parameter estimate of 0.12 for storage time is statistically significant but is substantially smaller than that of the fourth ranked input. The implication is that storage time by itself contributes little to variation in dose. The 95 percent confidence intervals for all parameter estimates are non-overlapping.

In comparing results for different food categories, it is apparent that the initial LM concentration is the top ranked input in all five food categories. The parameter estimates for initial LM concentration are larger by a factor of 1.1 to 2.9 than for the second largest inputs. The second most important input varies from one food category to another. For milk, pâtés and smoked seafood storage temperature was ranked second with parameter estimates a factor of 1.1, 1.7 and 2.9 smaller than the top ranked input. For deli salad and fresh soft cheese

growth at 5 $^0$C was ranked second with parameter estimates a factor of 2.1 and 3.0 smaller than the top ranked input. Serving size is ranked last three of the five times and storage time and storage temperature in other two cases.

In most cases the 95 percent confidence intervals are mutually exclusive and hence the rankings are robust. The exceptions are that the 95 percent confidence interval for storage time and growth at 5 $^0$C overlap for pâtés and serving size and storage time overlap for smoked seafood. Serving size and storage time are typically the two least important inputs, as they are never ranked higher than third. All of the parameter estimates for each of five inputs for all food categories are statistically significant.

The rankings based upon the analysis of co-mingled variability and uncertainty differs slightly from the rankings based upon the analysis of variability only as given in the previous section. For example, for pâtés, the initial LM concentration was ranked first in the former case but second in the latter case. Although the parameter estimates are significantly different from each other for the first and second ranked inputs in both cases, the magnitudes of the parameter estimates for each input do not differ substantially. The parameter estimate for the initial LM concentration is a factor of 1.1 higher than that for the storage temperature in the former case whereas the parameter estimate for the storage temperature is a factor of 1.1 higher than that for the initial LM concentration in the latter case. Thus, these parameter estimates differ by only approximately 10 percent. This difference is likely attributable to differences in the range of values and the relatively likelihood of values for these two inputs when comparing the two different simulations.

The $R^2$ value is above 0.80 for all food categories as in the case of the variability only dataset. Hence the linearity assumption holds well for the models.

### 15.1.3 Sensitivity Analysis of Exposure Module Based Upon Two Dimensional Simulation of Variability and Uncertainty

This section covers application of regression analysis to the exposure module when variability was considered under various uncertainty realizations. Thus, the variability and uncertainty dimensions were clearly distinguished. Regression analysis was performed for each of five selected food categories. The dataset was generated from a two dimensional Monte Carlo simulation that had 250 uncertain realizations and 3000 variability iteration. The details of dataset generation are described in Section 12.12.

Regression analysis was performed separately for each of the 250 uncertainty iterations. The results of these analyses are summarized in Tables 15-3 and 15-4. Table 15-3 displays the mean and median value of the parameter estimate over the 250 iterations, the 95 percent probability range of the coefficient, the frequency of statistically significant parameter estimate, the mean and median rank, and the 95% probability range of the rank. Table 15-4 provides additional information regarding the frequency over the 250 iterations with which each input was assigned a particular rank. Furthermore, the results are illustrated graphically for each food group in Figures 15-1 through 15-5 for deli salad, fresh soft cheese, milk, pate, and smoked seafood, respectively.

The interpretation of the results is explained using deli salad as the primary example. In this case, in each of the 250 iterations of uncertainty, the initial LM concentration was identified as the most important of the inputs with respect to variability. Thus, although there was uncertainty in the model inputs, there is no ambiguity in this case regarding the top ranked input. In the previous analyses for variability only and both variability and uncertainty co-mingled, the initial LM concentration was also identified as the top ranked input. However, the results for the second ranked input have some ambiguity. Growth at 5 $^0$C was identified as the second ranked input for 70 percent of the uncertainty iterations, as well as in the previous two case studies. Serving size was ranked second for 16 percent of the iterations and storage temperature was ranked second for 13 percent of the iterations. Thus, there is a chance that any of these three inputs may be the second most important input. Furthermore, it is more probable although not certain that the growth at 5 $^0$C is the second most important input.

The three inputs of serving size, storage temperature, and growth at 5 $^0$C appear to vie for the second, third, and fourth rankings. Each of these three inputs has a probability ranging from as low as 10 percent to as high as 70 percent of having one of these three ranks. The most likely outcome is for the growth rate at 5 $^0$C to be ranked second, with 70 percent probability, the serving size to be ranked third, with 54 percent probability, and the storage temperature to be ranked fourth, with 61 percent probability. These most likely rankings are also consistent with the rankings of the previous two case studies.

The higher the probability of a particular input having a particular rank and the narrower the range of possible ranks for a given input, the more robust is the ranking of that input. For example, the first and fifth ranked inputs have a 100 percent probability of their respective

Table 15-3. Parameter Estimates and Within Food Category Rank Statistics for Variability Under Uncertainty for the *Listeria monocytogenes* Exposure Model Inputs for Deli Salad, Fresh Soft Cheese, Milk, Pâtés, and Smoked Seafood

| Variable Name | Parameter Estimate | | | Rank | |
|---|---|---|---|---|---|
| | Mean, Median | 95% C.I. | Freq | Mean, Median | Range |
| **Deli Salad** | | | | | |
| Serving Size (g) | 0.34, 0.32 | 0.24, 0.47 | 249 | 3.1,3 | 2-4 |
| Initial LM Concentration (log cfu/g) | 0.76,0.76 | 0.62, 0.87 | 250 | 1,1 | 1-1 |
| Storage temperature ($^0$C) | -0.29, -0.31 | -0.43, -0.11 | 250 | 3.5,4 | 2-4 |
| Storage time (days) | -0.14, -0.13 | -0.22, -0.07 | 250 | 5,5 | 5-5 |
| Growth at 5 $^0$C (log cfu/day) | 0.40, 0.40 | 0.30, 0.52 | 250 | 2.4,2 | 2-4 |
| **Fresh Soft Cheese** | | | | | |
| Serving Size (g) | 0.35, 0.34 | 0.21, 0.52 | 250 | 2.5,2 | 2-3 |
| Initial LM Concentration (log cfu/g) | 0.91, 0.93 | 0.71, 1.0 | 250 | 1,1 | 1-1 |
| Storage temperature ($^0$C) | 0.18, 0.17 | 0.10, 0.31 | 249 | 4,4 | 4-4 |
| Storage time (days) | 0.12, 0.11 | 0.05, 0.24 | 230 | 5,5 | 5-5 |
| Growth at 5 $^0$C (log cfu/day) | 0.35, 0.33 | 0.20, 0.51 | 250 | 2.5,3 | 2-3 |
| **Milk** | | | | | |
| Serving Size (g) | 0.33, 0.37 | 0.15, 0.44 | 250 | 3.5, 3 | 3-5 |
| Initial LM Concentration (log cfu/g) | 0.73, 0.73 | 0.51, 0.91 | 250 | 1.2, 1 | 1-2 |
| Storage temperature ($^0$C) | 0.58, 0.58 | 0.44, 0.73 | 250 | 1.8, 2 | 1-2 |
| Storage time (days) | 0.29, 0.29 | 0.18, 0.38 | 250 | 4.0, 4 | 3-5 |
| Growth at 5 $^0$C (log cfu/day) | 0.23, 0.24 | 0.12, 0.32 | 250 | 4.5, 5 | 3-5 |
| **Pâtés** | | | | | |
| Serving Size (g) | 0.20, 0.22 | 0.08, 0.27 | 250 | 4.6, 5 | 4-5 |
| Initial LM Concentration (log cfu/g) | 0.66, 0.65 | 0.53, 0.80 | 250 | 1.5, 1 | 1-2 |
| Storage temperature ($^0$C) | 0.65, 0.64 | 0.52, 0.81 | 250 | 1.5, 2 | 1-2 |
| Storage time (days) | 0.30, 0.29 | 0.17, 0.44 | 250 | 3.6, 4 | 3-5 |
| Growth at 5 $^0$C (log cfu/day) | 0.31, 0.34 | 0.16, 0.47 | 250 | 3.7, 3 | 3-5 |
| **Smoked Seafood** | | | | | |
| Serving Size (g) | 0.20, 0.21 | 0.12, 0.26 | 250 | 3.3, 3 | 3-4 |
| Initial LM Concentration (log cfu/g) | 0.87, 0.89 | 0.50, 0.96 | 250 | 1.0, 1 | 1-2 |
| Storage temperature ($^0$C) | 0.38, 0.37 | 0.24, 0.60 | 250 | 2.0, 2 | 1-2 |
| Storage time (days) | 0.17, 0.17 | 0.08, 0.27 | 250 | 3.7, 4 | 3-4 |
| Growth at 5 $^0$C (log cfu/day) | 0.12, 0.12 | 0.06, 0.17 | 250 | 5.0, 5 | 5-5 |

Table 15-4. Probability of Within Food Category Rankings for Variability Under Uncertainty for the *Listeria monocytogenes* Exposure Model Inputs for Deli Salad, Fresh Soft Cheese, Milk, Pâtés, and Smoked Seafood

| Variable Name | Rank 1 | Rank 2 | Rank 3 | Rank 4 | Rank 5 |
|---|---|---|---|---|---|
| **Deli Salad** | | | | | |
| Serving Size (g) | | 0.16 | 0.54 | 0.29 | |
| Initial LM Concentration (log cfu/g) | 1.00 | | | | |
| Storage temperature ($^0$C) | | 0.13 | 0.26 | 0.61 | |
| Storage time (days) | | | | | 1.00 |
| Growth at 5 $^0$C (log cfu/day) | | 0.70 | 0.20 | 0.10 | |
| **Fresh Soft Cheese** | | | | | |
| Serving Size (g) | | 0.52 | 0.48 | | |
| Initial LM Concentration (log cfu/g) | 1.00 | | | | |
| Storage temperature ($^0$C) | | | | 1.00 | |
| Storage time (days) | | | | | 1.00 |
| Growth at 5 $^0$C (log cfu/day) | | 0.48 | 0.52 | | |
| **Milk** | | | | | |
| Serving Size (g) | | | 0.69 | 0.14 | 0.17 |
| Initial LM Concentration (log cfu/g) | 0.79 | 0.21 | | | |
| Storage temperature ($^0$C) | 0.21 | 0.79 | | | |
| Storage time (days) | | | 0.19 | 0.64 | 0.17 |
| Growth at 5 $^0$C (log cfu/day) | | | 0.12 | 0.22 | 0.66 |
| **Pâtés** | | | | | |
| Serving Size (g) | | | 0.01 | 0.33 | 0.66 |
| Initial LM Concentration (log cfu/g) | 0.51 | 0.49 | | | |
| Storage temperature ($^0$C) | 0.50 | 0.50 | | | |
| Storage time (days) | | | 0.48 | 0.42 | 0.10 |
| Growth at 5 $^0$C (log cfu/day) | | | 0.51 | 0.25 | 0.24 |
| **Smoked Seafood** | | | | | |
| Serving Size (g) | | | 0.66 | 0.34 | |
| Initial LM Concentration (log cfu/g) | 0.97 | 0.03 | | | |
| Storage temperature ($^0$C) | 0.03 | 0.97 | | | |
| Storage time (days) | | | 0.34 | 0.64 | 0.02 |
| Growth at 5 $^0$C (log cfu/day) | | | | 0.02 | 0.98 |

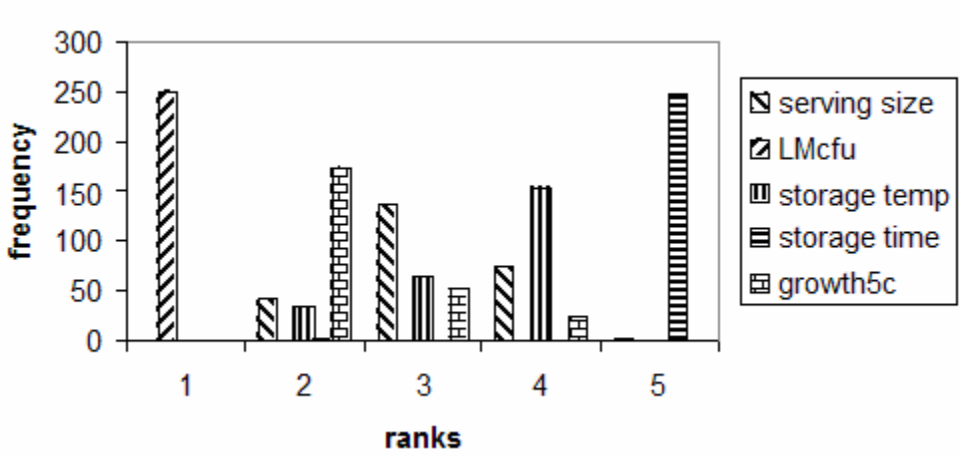



Figure 15-1. Frequency of Input Ranks for the Deli Salad Food Category in the Exposure Module of The *Listeria monocytogenes* Model.

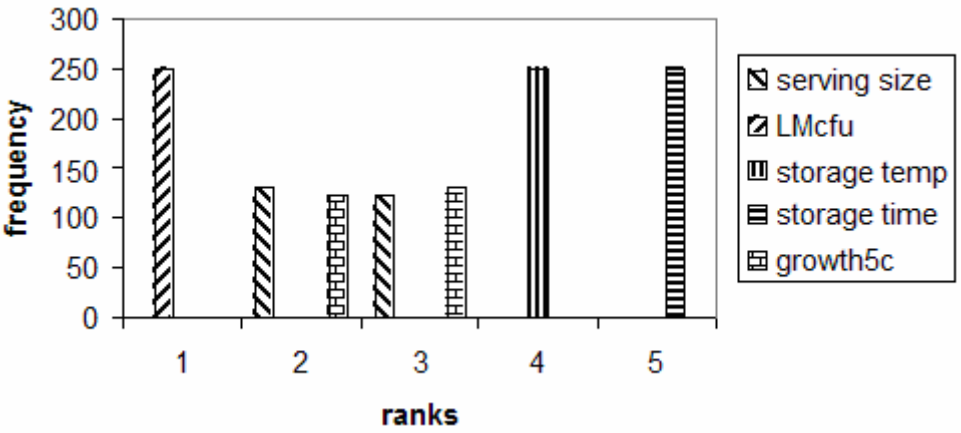



Figure 15-2. Frequency of Input Ranks for the Fresh Soft Cheese Food Category in the Exposure Module of The *Listeria monocytogenes* Model.

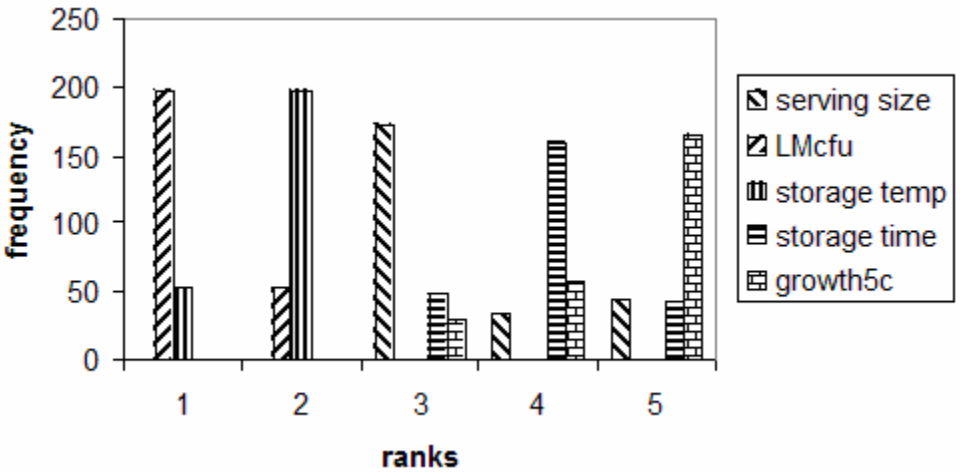



Figure 15-3. Frequency of Input Ranks for the Milk Food Category in the Exposure Module of The *Listeria monocytogenes* Model.

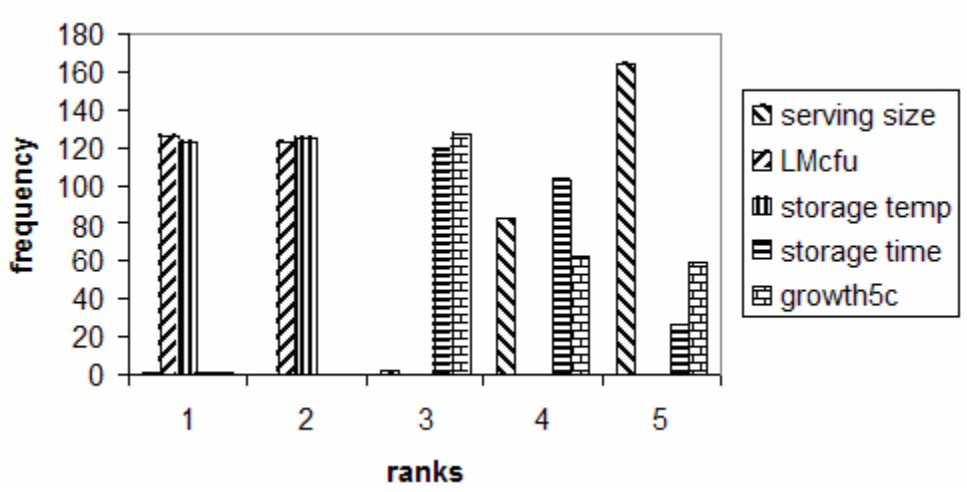



Figure 15-4. Frequency of Input Ranks for the Pâtés Food Category in the Exposure Module of The *Listeria monocytogenes* Model.

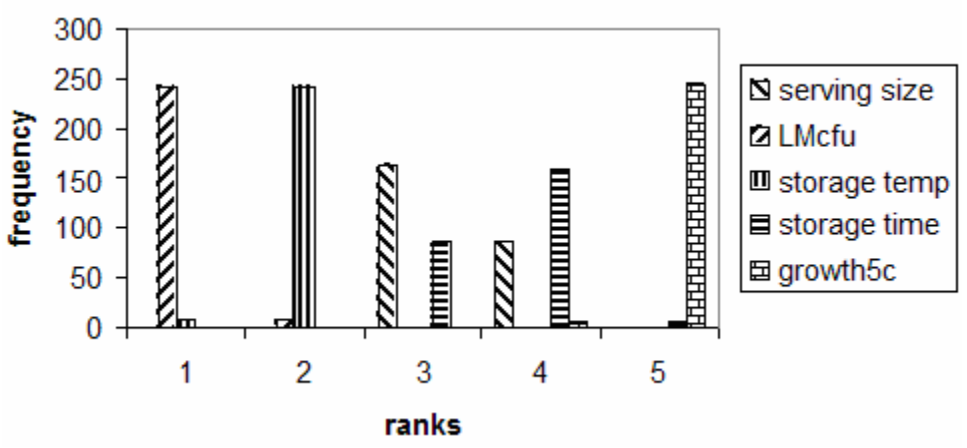



Figure 15-5. Frequency of Input Ranks for the Smoked Sea Food Category in the Exposure Module of The *Listeria monocytogenes* Model.

rankings. In contrast, the second through fourth ranked inputs could take on any of these three ranks. Thus, despite the uncertainty, it is possible to make a certain determination that the initial LM concentration is the most important input and that the storage time is the least important input. Because of uncertainty, there is some ambiguity regarding the ranks of the other three inputs, although for each of these three there is clearly a most probable rank that can be assigned.

The results for fresh soft cheese are less ambiguous than those for deli salad. There is a 100 percent probability associated with the first, fourth, and fifth ranked inputs. By comparison, there is nearly an equal chance that the second and third ranked inputs could reverse position in

the rank order. These results suggest that the inputs could be ranked as follows: (1) a clear top ranked input; (2) a group of two inputs that are of equal importance and that are less important than the top ranked input but more important than the last two ranked inputs; and (3) two inputs of ranks four and five that are clearly less important than any others. Thus, it is possible to discriminate among individual inputs or groups of inputs.

The results for milk and pâtés illustrate situations in which there is considerable ambiguity in the rank ordering of all of the inputs. Nonetheless, even when uncertainty leads to ambiguity in the rankings, there are some inferences that can be made with confidence. For example, although there is ambiguity regarding whether the initial LM concentration or the storage temperature should be ranked first for milk, it is clear that these two inputs are ranked either first or second and, therefore, are more important than any of the other three inputs. Moreover, although each of the other three inputs has a probability of taking on any rank between three and five, there is a clear most likely rank for each of these. For example, serving size has a 69 percent probability of being ranked third, storage time has a 64 percent probability of being ranked fourth, and growth at 5 $^0$C has a 66 percent probability of being ranked fifth.

For pâtés, the initial LM concentration and storage temperature have almost an equal probability of being ranked first or second, but neither is ranked lower than second in any of the 250 uncertainty iterations. The storage time and growth at 5 $^0$C have substantial probabilities of being ranked third or fourth. The serving size has a high probability of being ranked fifth and is unlikely to be ranked as high as third. Thus, it is possible to identify groups of inputs that have differing importance. In this case, the initial LM concentration and storage temperature are of highest importance. The storage time and growth rate at 5 $^0$C are of secondary importance. The serving size is probably the least important input.

The results for smoked seafood have a nearly 100 percent probability associated with the first, second, and fifth ranked inputs. However, there is ambiguity in ranking of third and fourth most important input. In a few cases the top two inputs and the least important two inputs interchange ranks. The third and fourth ranked inputs of serving size and storage time have ambiguity in ranking. However, the likelihood of serving size being ranked third is almost twice that compared to storage time.

In summary, the uncertainty analysis presented here regarding identification of the most important factors that influence variability in the exposure model indicate that although there can

be ambiguity regarding the ranking of inputs, it is often possible to develop robust insight into the rank to assign to a given input or to a group of inputs that have similar importance. Thus, uncertainty analysis of the sensitivity analysis results is a means to gain confidence regarding the robustness of the rankings, and to raise awareness of situations in which two or more inputs may be of comparable importance.

## 15.2 Application of Regression Analysis to the Dose Response Module

Regression analysis was applied to the dose response module for five selected food groups, including deli salad, fresh soft cheese, milk, pâtés, and smoked seafood. For each food group, the inputs of interest include uncertainty in dose adjustment factor, exposure period, virulence susceptibility uncertainty, and mouse lethality uncertainty and population exposure. The output of interest is mortality ascertained to the food group for the neonatal sub-population.

The dose response module incorporates only uncertainty. This section focuses on application of regression analysis based upon uncertainty in inputs. The results of regression analysis for each input and for the significance of the statistical model are shown in Table 15-5. For each food category, the five dose response models inputs are assigned a sensitivity rank based upon the magnitude of their respective parameter estimates. For example, for deli salad, the fraction of population exposed has a parameter estimate of 0.07, which is a factor of 1.4 greater than the input with the second highest parameter estimate. Furthermore, the parameter estimate for the fraction of population exposed is statistically significant. Thus, fraction of population exposed is inferred to be the most sensitive input for this food category, and it is more sensitive than the second ranked input of dose adjustment factor.

The magnitudes of the parameter estimates for the second and third ranked inputs differ by a factor of 1.4. However, the confidence intervals for the magnitudes of these parameter estimates overlap substantially, implying that there is ambiguity regarding the rank order among these two. Furthermore, although the magnitude of the parameter estimates for the top ranked input is factors of 1.7 greater than that for the third ranked input, the magnitude of the confidence intervals for these two also overlap substantially. Therefore, there is ambiguity regarding the rank to assign to each of the top three inputs. However, because these three inputs are statistically significant, and the remaining two are not, it is clear that the top three are more important than the other two. Furthermore, the two inputs that are not statistically significant have no substantial impact on the output.

Table 15-5. Parameter Estimates and Within Food Category Rankings for Uncertainty Only for the *Listeria monocytogenes* Dose Response Model Inputs for Deli Salad, Fresh Soft Cheese, Milk, Pâtés, and Smoked Seafood

| Variable Name | Parameter Estimate | 95% Confidence Interval | Rank Within The Food Category | Statistically Significant |
|---|---|---|---|---|
| **Deli Salad ($R^2$ = 0.01)** | | | | |
| Dose Adjustment Factor | -0.048 | -0.080, -0.015 | 2 | Yes |
| Exposure Period | 0.038 | 0.006, 0.069 | 3 | Yes |
| Virulence Susceptibility | -0.014 | -0.045, 0.017 | NA | No |
| Mouse Lethality | -0.006 | -0.038, 0.025 | NA | No |
| Fraction of Population Exposed | -0.065 | -0.096, -0.034 | 1 | Yes |
| **Fresh Soft Cheese ($R^2$ = 0.01)** | | | | |
| Dose Adjustment Factor | -0.088 | -0.120, -0.056 | 1 | Yes |
| Exposure Period | 0.013 | -0.019, 0.045 | NA | No |
| Virulence Susceptibility | -0.018 | -0.049, 0.013 | NA | No |
| Mouse Lethality | 0.006 | -0.025, 0.037 | NA | No |
| Fraction of Population Exposed | -0.025 | -0.056, 0.006 | NA | No |
| **Milk ($R^2$ = 0.01)** | | | | |
| Dose Adjustment Factor | -0.107 | -0.139, -0.075 | 1 | Yes |
| Exposure Period | 0.045 | 0.013, 0.077 | 2 | Yes |
| Virulence Susceptibility | -0.005 | -0.036, 0.026 | NA | No |
| Mouse Lethality | -0.001 | -0.032, 0.031 | NA | No |
| Fraction of Population Exposed | 0.028 | -0.003, 0.059 | NA | No |

(Continued on next page.)

Table 15-5. Continued

| Variable Name | Parameter Estimate | 95% Confidence Interval | Rank Within The Food Category | Statistically Significant |
|---|---|---|---|---|
| **Pâtés($R^2$ = 0.02)** | | | | |
| Dose Adjustment Factor | -0.146 | -0.178, -0.114 | 1 | Yes |
| Exposure Period | 0.050 | 0.019, 0.082 | 2 | Yes |
| Virulence Susceptibility | -0.028 | -0.059, 0.003 | NA | No |
| Mouse Lethality | -0.007 | -0.038, 0.024 | NA | No |
| Fraction of Population Exposed | 0.018 | -0.013, 0.048 | NA | No |
| **Smoked Seafood ($R^2$ = 0.02)** | | | | |
| Dose Adjustment Factor | -0.091 | -0.123, -0.059 | 1 | Yes |
| Exposure Period | 0.035 | 0.004, 0.067 | 2 | Yes |
| Virulence Susceptibility | 0.012 | -0.019, 0.043 | NA | No |
| Mouse Lethality | -0.0164 | -0.048, 0.015 | NA | No |
| Fraction of Population Exposed | 0.030 | -0.001, 0.061 | NA | No |

For three of the food groups, including milk, pâtés, and smoked seafood, the dose adjustment factor and the exposure period are the two most important inputs. The other three inputs are not statistically significant and therefore are judged to be unimportant. For fresh soft cheese, the only two statistically significant inputs are the dose adjustment factor and the fraction of population exposed. Thus, for all food groups, it is possible to clearly identify inputs that are unimportant as distinguished from those to which the output is sensitive. In some cases, the statistically significant inputs can be distinguished from each other. For example, the top ranked input for pâtés is significantly more important than the second ranked input because the confidence intervals for the parameter estimates do not overlap.

The $R^2$ value is very low for all the food categories. The linear model assumption is not valid here. Hence the parameter estimates are not reliable.

## 15.3 Summary of the Result of Application of Regression Analysis

In Sections 15-1 to 15-2 regression analysis was applied to different modules of the *Listeria monocytogenes* model. Regression analysis was evaluated based upon applicability of

the functional form of the model, the use of regression coefficients as an indicator of sensitivity, the use of confidence intervals for regression coefficients to evaluate the robustness of rankings, and the ease of application.

The need to assume a specific functional relation between inputs and the output in a regression model can be considered as a disadvantage for this method. If the specific functional assumption is not valid compared to the original model, the results from the regression analysis also may not be valid. In these cases the fitted regression model addresses only a portion of the original model response variation. Although a better fit to the data may be obtained by use of different basis functions, the selection of appropriate basis functions is often by trial and error. Furthermore, sensitivity analysis methods that do not require specification of a specific functional relation are expected to be favored for models with non-linear responses. ANOVA and CART are examples of these kinds of sensitivity analysis methods. The application of these methods to the *Listeria monocytogenes* model is presented in Chapters 16 and 17, respectively.

The estimated $R^2$ for the fitted linear models to the exposure module of the *Listeria monocytogenes* model indicated that the linear assumption for the model response was a good approximation in both variability only and co-mingled variability and uncertainty cases. The $R^2$ values ranged from 0.80 to 0.92. However, the linearity assumption in dose response module appears to be invalid since the $R^2$ values were exceptionally small. Methods such as rank regression, as demonstrated in Section 2.2.3, may be applied to gain insight regarding monotonic response if any. However, the response may be highly random and thus it is possible that no form of regression will produce a good fit.

The use of regression coefficient estimates as a measure of sensitivity of the output to individual inputs was demonstrated in this chapter. However, these coefficients do not directly enable insight regarding their precision and accuracy. Each regression coefficient has a standard error that can be used to derive the confidence interval for the regression coefficient in order to evaluate the robustness of the coefficient.

The application of regression analysis provides useful parameters via which the relative importance of the inputs can be ranked. In the case of two-dimensional analysis where variability was considered under different uncertain realizations, a thorough understanding of the rankings for the inputs was achieved. The variability only case gave robust ranking for inputs that were either highly important or not important. For example, in the case of smoked seafood, the

variability only case identified initial LM concentration and storage temperature as top two important inputs and growth at 5 $^{0}$C as the least important input. In some cases, the results from the two dimensional simulation imply more ambiguity regarding the rankings than do the results from a one-dimensional simulation of variability only. For example, the ranking for key sources of variability in exposure from consumption of pâté were more ambiguous in the two-dimensional case than in the one-dimensional case. However, both approaches produced similar insights regard the relative importance of various groups of inputs.

With regard to sensitivity analysis, it may be the case in some situations that a one dimensional simulation of variability and uncertainty may be adequate as a basis for identifying important inputs, as suggested by the results in this chapter. However, for policy and decision making purposes, a two-dimensional simulation in which variability and uncertainty are distinguished may be required. Thus, during model development, it may be reasonable to start with a one dimensional analysis for the purpose of identifying priorities for development of ranges and distributions of model inputs. However, the fact that the rankings based upon sensitivity analysis were similar in the one and two-dimensional probabilistic simulations of these case studies does not imply that such similarities will always be found in other case studies and with other models.

# 16 APPLICATION OF ANALYSIS OF VARIANCE TO THE *LISTERIA MONOCYTOGENES* MODEL

The purpose of this chapter is to apply ANOVA to the *Listeria monocytogenes* model. ANOVA is discussed in Section 2.2.3. The *Listeria monocytogenes* model is discussed in Chapter 12. ANOVA assumes a generalized functional form and hence it is applicable to both linear and non-linear models. The *Listeria monocytogenes* model is nonlinear. Therefore, a key consideration in this chapter is to evaluate whether the non-linearity in *Listeria monocytogenes* model is adequately captured by ANOVA.

The input variables in each module are treated as "factors" for ANOVA as described in Section 2.2.3. The outputs of each of the modules are referred to as "response" variables. Since there are multiple factors, multifactor ANOVA is applicable as described in Section 2.2.3. All input variables of interest in the *Listeria monocytogenes* model are continuous variables. ANOVA assumes that variables have discrete levels. Thus, to allow application of ANOVA, continuous distributions are partitioned into discrete ranges to create levels as described in Section 12.9. A "treatment" is a combination of a level of each of the factors.

ANOVA uses the F test to determine whether a factor has a significant effect on the mean treatment response over the combination of other factors. If the null hypothesis, which is no difference among treatment means for different levels of a factor, is accepted, then the implication is that a change in levels of the factor does not change the response significantly. When the null hypothesis is rejected, then the implication is that changing the level of the factor changes the response significantly. In the latter case, more analysis is done to understand the exact effects of each of the levels of the factor on the response.

ANOVA allows for understanding of any interaction between factors in the model. For example, for factors with significant interactions, various contrasts can be tested. As an example a question that can be answered using ANOVA is: when storage temperature is high, is there a significant difference between the responses when values of storage time are changed from a high level to a lower level.

In order to apply ANOVA to different modules of the *Listeria monocytogenes* food safety risk assessment model, SAS© Version 8.02 was used. This software package has the ability to perform ANOVA by using available procedures on a dataset. The macro feature in SAS allows programs to be written to automate these procedures for application to multiple datasets.

This chapter is divided into three parts. The results of application of ANOVA to the exposure module are presented in Section 16.1. In this section, datasets obtained from simulations considering only variability, and both variability and uncertainty are analyzed. The relative rankings of the input variables and contrasts for storage time and temperature interaction are presented for each of five selected food categories. Section 16.2 presents the results of application of ANOVA on the dose response module. In this section a dataset from an uncertainty simulation is analyzed for all five selected food categories. The key findings and implications based upon ANOVA are discussed in Section 16.3.

## 16.1 Application of ANOVA to the Exposure Module

ANOVA was applied to the exposure module for five selected food groups, including deli salad, fresh soft cheese, milk, pâtés, and smoked seafood. For each food group, the inputs of interest include serving size in grams, initial LM concentration in log cfu/g, storage temperature in $^0$C, storage time in days, and growth potential at 5 $^0$C in log cfu/day. The output of interest is the dose value corresponding to each meal serving size simulated. Thus for ANOVA application, the factors in the exposure module are the five inputs and the response variable is the output.

In order to perform ANOVA, each factor should be divided into discrete levels. For continuous inputs, levels were defined by splitting the domain of values into ranges based upon the cumulative distribution function of the input. In particular, levels were defined based upon the lower tail, middle region, and upper tail of the distribution of each input. For each input, Level 1 was defined to include the values of the input up to the 20$^{th}$ percentile of the distribution. Level 2 was defined to include values between the 20$^{th}$ percentile and 80$^{th}$ percentile. Level 3 was defined to include values above the 80$^{th}$ percentile. In addition to performing ANOVA based upon the individual contribution of each of the factors, the possibility of important interactions among the factors was considered. In particular, analysts and others have typically expressed interest in the interaction between storage time and storage temperature with respect to the dose. Therefore, the interaction between these two factors was explicitly considered. In addition, when an interaction effect is statistically significant, a more detailed technique for evaluating the importance of inputs is to consider contrasts. Contrasts test whether there is a significant difference in mean responses between any two given levels of a factor, such as storage time, when another factor, such as storage temperature, is kept at a specific level. The

change in mean response is evaluated based upon contrasts of storage time between high and low levels, high and medium levels, and medium and low levels.

The following subsection focuses on application of ANOVA based upon only variability in inputs and the corresponding estimated variability in dose. Section 16.1.2 focuses upon application of ANOVA when both variability and uncertainty are considered in a one-dimensional simulation. The purpose of conducting both analyses is to gain insight into how the importance of each factor may differ depending on the distributions assigned to each input.

### 16.1.1 Sensitivity Analysis of the Exposure Model Based Upon Variability Only

This section covers application of ANOVA to the exposure module when only the variability dimension was considered during simulation. Since the original exposure module considers both variability and uncertainty, the exposure module was modified to remove the uncertainty dimension. The uncertainty dimension of variables was fixed to median values as described in Section 12.9.

ANOVA was performed for each of five selected food categories. The assumptions regarding factor levels for the inputs for the five food categories are given in Table 16-1. The results of ANOVA for each individual factor, for the interaction of storage temperature and storage time, and for the significance of the statistical model are shown in Table 16-2.

For each food category, the five factors are assigned a sensitivity rank based upon the magnitude of their respective F-values. For example, for deli salad, the initial LM concentration has an F-value of 1,540, which is a ratio of 4.0 greater than the input with the second highest F-value. Furthermore, the F-value for the initial LM concentration is statistically significant. Thus, the initial LM concentration is inferred to be the most sensitive input for this food category, and it is substantially more sensitive than the second ranked input of growth rate at 5 $^0$C. The F-value of the second ranked input is a ratio of 1.6 greater than that for the third ranked input. Thus, the second and third ranked inputs have F-values that do not differ substantially. Therefore, these two inputs may be of approximately comparable importance. The smallest F-value of 39 for storage time is statistically significant but is substantially smaller than that of the fourth ranked input. The implication is that storage time by itself contributes little to variation in dose.

In comparing results for different food categories, it is apparent that the initial LM concentration is the top ranked input in all cases and that the F-values for the initial LM

Table 16-1. Factor Levels for Variability Only for the *Listeria monocytogenes* Exposure Model Inputs for Deli Salad, Fresh Soft Cheese, Milk, Pâtés, and Smoked Seafood

| Factor | Level 1[(a)] | Level 2[(b)] | Level 3[(c)] |
|---|---|---|---|
| **Deli Salad** | | | |
| Serving Size (g) | 0 to 80 | 80 to 210 | 210 to 1420 |
| Initial LM Concentration (log cfu/g) | -3.35 to -2.17 | -2.17 to -0.88 | -0.88 to 3 |
| Storage temperature ($^0$C) | 0 to 4 | 4 to 5 | 5 to 22 |
| Storage time (days) | 0 to 2.6 | 2.6 to 5 | 5 to 9.7 |
| Growth at 5 $^0$C (log cfu/day) | -0.35 to -0.17 | -0.17 to 0 | 0 to 0.145 |
| **Fresh Soft Cheese** | | | |
| Serving Size (g) | 0 to 30 | 30 to 40 | 40 to 250 |
| Initial LM Concentration (log cfu/g) | -2.7 to -1.9 | -1.9 to 0 | 0 to 7.2 |
| Storage temperature ($^0$C) | 0 to 2.5 | 2.5 to 7 | 7 to 23 |
| Storage time (days) | 0.5 to 2.5 | 2.5 to 7 | 7 to 23 |
| Growth at 5 $^0$C (log cfu/day) | -0.5 to -0.05 | -0.05 to 0.2 | 0.2 to 0.7 |
| **Milk** | | | |
| Serving Size (g) | 0 to 150 | 150 to 350 | 350 to 3825 |
| Initial LM Concentration (log cfu/g) | -3.8 to -2.4 | -2.4 to -0.9 | -0.9 to 4.8 |
| Storage temperature ($^0$C) | 0 to 2.5 | 2.5 to 5.5 | 5.5 to 22 |
| Storage time (days) | 0.7 to 3 | 3 to 6 | 6 to 12 |
| Growth at 5 $^0$C (log cfu/day) | 0.03 to 0.16 | 0.16 to 0.35 | 0.35 to 0.91 |
| **Pâtés** | | | |
| Serving Size (g) | 0 to 50 | 50 to 110 | 110 to 455 |
| Initial LM Concentration (log cfu/g) | -3 to -1.8 | -1.8 to 1 | 1 to 9.1 |
| Storage temperature ($^0$C) | 0 to 2.5 | 2.5 to 5.5 | 5.5 to 22 |
| Storage time (days) | 0.5 to 5.5 | 5.5 to 13.5 | 13.5 to 29 |
| Growth at 5 $^0$C (log cfu/day) | 0.14 to 0.18 | 0.18 to 0.32 | 0.32 to 0.37 |
| **Smoked Seafood** | | | |
| Serving Size (g) | 0 to 60 | 60 to 85 | 85 to 142 |
| Initial LM Concentration (log cfu/g) | -2.5 to -1.5 | -1.5 to 1 | 1 to 9 |
| Storage temperature ($^0$C) | 0 to 2.5 | 2.5 to 6 | 6 to 22 |
| Storage time (days) | 0.5 to 3 | 3 to 8 | 8 to 23 |
| Growth at 5 $^0$C (log cfu/day) | 0.05 to 0.1 | 0.1 to 0.18 | 0.18 to 0.27 |

(a) Lower 0.2 percentile range; (b) 0.2 to 0.8 percentile range; (c) 0.8 to 1.0 percentile range

Table 16-2. ANOVA Results for Variability Only for the Main Effects of Individual Inputs, Interactions, and Statistical Model Significance for the *Listeria monocytogenes* Exposure Model for Deli Salad, Fresh Soft Cheese, Milk, Pâtés, and Smoked Seafood

| Factor | F Value | Pr > F | Significant | Rank |
|---|---|---|---|---|
| **Deli Salad** | | | | |
| Serving Size (g) | 236 | <.0001 | Yes | 3 |
| Initial LM Concentration (log cfu/g) | 1540 | <.0001 | Yes | 1 |
| Storage temperature ($^0$C) | 142 | <.0001 | Yes | 4 |
| Storage time (days) | 39 | <.0001 | Yes | 5 |
| Growth at 5 $^0$C (log cfu/day) | 383 | <.0001 | Yes | 2 |
| Interaction (temp*time) | 13 | <.0001 | Yes | |
| Model | 168 | <.0001 | Yes | |
| **Fresh Soft Cheese** | | | | |
| Serving Size (g) | 105 | <.0001 | Yes | 2 |
| Initial LM Concentration (log cfu/g) | 1189 | <.0001 | Yes | 1 |
| Storage temperature ($^0$C) | 5 | 0.0112 | Yes | 4 |
| Storage time (days) | 0 | 0.6301 | No | 5 |
| Growth at 5 $^0$C (log cfu/day) | 80 | <.0001 | Yes | 3 |
| Interaction (temp*time) | 9 | <.0001 | Yes | |
| Model | 175 | <.0001 | Yes | |
| **Milk** | | | | |
| Serving Size (g) | 74 | <.0001 | Yes | 4 |
| Initial LM Concentration (log cfu/g) | 675 | <.0001 | Yes | 1 |
| Storage temperature ($^0$C) | 269 | <.0001 | Yes | 2 |
| Storage time (days) | 62 | <.0001 | Yes | 5 |
| Growth at 5 $^0$C (log cfu/day) | 96 | <.0001 | Yes | 3 |
| Interaction (temp*time) | 22 | <.0001 | Yes | |
| Model | 89 | <.0001 | Yes | |
| **Pâtés** | | | | |
| Serving Size (g) | 7 | 0.0014 | Yes | 5 |
| Initial LM Concentration (log cfu/g) | 658 | <.0001 | Yes | 1 |
| Storage temperature ($^0$C) | 250 | <.0001 | Yes | 2 |
| Storage time (days) | 70 | <.0001 | Yes | 3 |
| Growth at 5 $^0$C (log cfu/day) | 49 | <.0001 | Yes | 4 |
| Interaction (temp*time) | 26 | <.0001 | Yes | |
| Model | 133 | <.0001 | Yes | |

(Continued on next page.)

Table 16-2. Continued

| Factor | F Value | Pr > F | Significant | Rank |
|---|---|---|---|---|
| **Smoked Seafood** | | | | |
| Serving Size (g) | 77 | <.0001 | Yes | 5 |
| Initial LM Concentration (log cfu/g) | 5074 | <.0001 | Yes | 1 |
| Storage temperature ($^0$C) | 535 | <.0001 | Yes | 2 |
| Storage time (days) | 163 | <.0001 | Yes | 3 |
| Growth at 5 $^0$C (log cfu/day) | 108 | <.0001 | Yes | 4 |
| Interaction (temp*time) | 65 | <.0001 | Yes | |
| Model | 440 | <.0001 | Yes | |

concentration are substantially larger by a ratio of 2.5 to 12.1 than for the second most important inputs. The second most important input varies from one food category to another. For milk, pate, and smoked seafood, the storage temperature is the second most important factor. Although the F-value for storage temperature in these three cases is substantially less than that for initial LM concentration, it is a ratio of 2.8 to 3.5 greater than that of the third ranked input. Thus, the relative importance of storage temperature for these three foods is unambiguous. For fresh soft cheese the serving size is identified as the second most important input. However, its F-value is only modestly larger than that for the third ranked input of growth rate at 5 $^0$C. Thus, both serving size and growth rate at 5 $^0$C are considered to be of comparable importance.

In many cases, it is possible to clearly identify the least important inputs and to do so without any ambiguity. The clearest example is for storage time for fresh soft cheese, which has an F-value that is not statistically significant. This indicates that there is no measurable relationship between the mean of the response and the different levels of this input. In other cases, the F-values are statistically significant but are small compared to the F-values of the highly ranked inputs. For example, for deli salad, the fifth ranked input of storage time has an F-value of 39 compared to an F-value of 1,540 for the highest ranked input. Thus, although there is a statistically significant relationship between the mean responses for dose with respect to different levels of storage time, the mean response for dose is substantially more strongly influenced by other inputs. Although storage time is ranked as highly as third for pate and smoked seafood, it is ranked fifth for deli salad, fresh soft cheese, and milk. Furthermore, the F-values for this input range from as low as 0 for fresh soft cheese to as high as only 163 for smoked seafood, compared to F-values ranging from 658 to 5,070 for the top ranked inputs in

each food category. Thus, on a comparative basis, it is possible to clearly distinguish the importance of the top ranked inputs vis-à-vis inputs with substantially smaller F-values.

The F-values for the interaction between storage temperature and storage time are typically low, ranging from 9 for fresh soft cheese to 65 for smoked seafood, when compared to the F-values for the top ranked inputs. For three of the food categories, deli salad, fresh soft cheese, and milk, either storage temperature and/or storage time are among the lowest ranked inputs. Therefore, it is not surprising that an interaction involving a relatively unimportant input would itself be unimportant. For pate and smoked seafood, storage temperature and storage time are the second and third, respectively, most important inputs. However, even in these cases, the F-value for the interaction term is small compared to the F-value of the top ranked inputs or to the F-values of either storage temperature or storage time on an individual basis. These results suggest that the interaction term is not of substantial importance.

The conclusions based upon the information presented in Tables 16-1 and 16-2 are summarized here:

- For all five food groups, the initial LM concentration is by far the most important input.
- For deli salad, the growth rate at 5 $^0$C, serving size, and storage temperature are of decreasing importance, respectively. The interaction between storage time and storage temperature is relatively unimportant. High storage temperatures are conducive to higher doses regardless of the storage time.
- For fresh soft cheese, serving size and growth at 5 $^0$C are of secondary importance. Storage time and storage temperature are not important.
- For milk, storage temperature is clearly the second most important input. Growth rate at 5 $^0$C, serving size, and storage time are of approximately comparable importance. The interaction effect between storage temperature and storage time is weak, but is clear that the higher level of storage temperature is of more importance than medium or lower levels.
- For pate and smoked seafood, storage temperature is clearly the second most important input. Storage time is substantially less important than storage temperature. The interaction term between storage time and storage temperature was of relatively small importance based upon the multifactor ANOVA. Growth rate at 5 $^0$C is the fourth most

important input for both of these food categories. Serving size is the least important input.

Although the interaction between storage time and storage temperature appears to be of less importance than the individual effect of many of the inputs, a more detailed evaluation of the interaction between these two inputs may reveal combinations of levels for these two factors that lead to substantial changes in the mean response. Therefore the F values, statistical significance and estimates of the contrast means are calculated. The ANOVA results for contrasts of these two factors are presented in Table 16-3 for all five selected food categories. An estimate is the difference between the mean responses at two levels considered for contrast. For deli salad, -0.134 is the estimate of difference between mean responses at high and low time levels when temperature was at low level. For example, for deli salads, three different levels of storage temperature are considered. For each level of storage temperature, three contrasts involving storage time are shown. The contrast from high to low indicates the sensitivity of the mean response for dose when storage time changes from a high level to a low level while storage temperature is held constant at a particular level.

The results demonstrate that the change in the mean response when storage time changes from a high to a low level is typically larger when storage temperature is at a high level than when storage temperature is at a medium or low level. Specifically, for deli salad the estimate for high storage temperature level and contrast between high and low storage time is -0.603, which is much larger than the estimate for the same contrast in storage time for medium or low levels of storage temperature. Furthermore, the estimate for all three possible contrasts for storage time associated with a high level of storage temperature is higher than the estimates for medium or low levels of storage temperature. This result is biologically plausible since higher temperatures are associated with higher growth rates of *Listeria monocytogenes*. If the temperature is in the low or medium levels, the growth rate is slow enough that there is less sensitivity to storage time. Qualitatively similar results were obtained for milk, pâtés and smoked seafood. For fresh soft cheese most of the contrasts were insignificant. For each of these food categories, the highest estimates for contrasts were associated with high levels of storage temperature.

Table 16-3. ANOVA Results for Variability Only for Contrasts of Storage Time and Storage Temperature for the *Listeria monocytogenes* Exposure Model for Deli Salad, Fresh Soft Cheese, Milk, Pâtés, and Smoked Seafood

| TEMPERATURE LEVELS [(a)] | CHANGE IN TIME LEVELS | | |
|---|---|---|---|
| | High-Low | High-Medium | Medium-Low |
| **Deli Salad** | | | |
| Low | 8.22, Yes, -0.134 | 5.30, Yes, -0.081 | 1.61, No, -0.053 |
| Medium | 16.18, Yes, -0.250 | 12.22, Yes, -0.176 | 2.05, No, -0.074 |
| High | 99.14, Yes, -0.603 | 58.66, Yes, -0.419 | 19.50, Yes, -0.184 |
| **Fresh Soft Cheese** | | | |
| Low | 8.01, Yes, -0.396 | 1.05, No, -0.101 | 5.59, Yes, -0.296 |
| Medium | 0.62, No, -0.099 | 0.65, No, 0.075 | 2.66, No, -0.173 |
| High | 0.96, No, 0.244 | 0.10, No, 0.076 | 2.07, No, 0.168 |
| **Milk** | | | |
| Low | 1.68, No, 0.146 | 0.31, No, 0.046 | 0.91, No, 0.101 |
| Medium | 77.13, Yes, 0.776 | 31.09, Yes, 0.416 | 25.94, Yes, 0.360 |
| High | 143.25, Yes, 1.660 | 56.77, Yes, 0.986 | 70.79, Yes, 0.673 |
| **Pâtés** | | | |
| Low | 5.92, Yes, 0.321 | 1.87, No, 0.126 | 2.60, No, 0.195 |
| Medium | 163.37 , Yes, 1.385 | 49.68, Yes, 0.612 | 80.38 , Yes, 0.772 |
| High | 48.26, Yes, 1.963 | 6.31, Yes, 0.692 | 184.65, Yes, 1.271 |
| **Smoked Seafood** | | | |
| Low | 1.87, No, 0.087 | 2.28, No, 0.069 | 0.09, No, 0.018 |
| Medium | 218.88, Yes, 0.666 | 99.27, Yes, 0.3735 | 65.48, Yes, 0.293 |
| High | 292.19, Yes, 1.725 | 91.84, Yes, 0.931 | 247.69, Yes, 0.795 |

(a) For Low, Medium and High levels of temperature the F value, statistical significance and estimate of difference between mean responses for each contrast are shown.

Conclusions regarding each temperature level corresponding to the three contrasts, for each of the five food categories are summarized here:

For deli salad,

- At low and medium temperature, exposure levels are sensitive only to the high levels of storage time.
- At high temperature, if storage time changes from one level to another the response will change significantly.

For fresh soft cheese,

- At low temperatures, it is important to constrain the storage time to the low level otherwise there is a significant change in response.
- At medium or high temperatures, controlling the storage time does not yield any control over response.

For milk,

- At low temperatures, the storage time level does not matter.
- At medium and high temperatures, any control measure that changes the storage time from one level to another will result in a significant change in response.

For pâtés,

- At low temperatures, only when storage time changes from a high to low level will there be a significant change in response.
- At medium and high temperatures, any control measure that changes the storage time at retail from one level to another will yield a significant change in response.

For smoked seafood,

- At low temperatures, the storage time level does not matter.
- At medium and high temperatures, any control measure that changes the storage time at retail from one level to another will yield a significant change in response.

### 16.1.2 Sensitivity Analysis of the Exposure Model Based Upon Co-mingled Variability and Uncertainty

This section covers application of ANOVA to the exposure module when both variability and uncertainty were considered in a one-dimensional simulation.

ANOVA was performed for each of five selected food categories. The input assumptions regarding factor levels for the inputs for the five food categories are given in Table 16-4. The

results of ANOVA for each individual factor, for the interaction of storage temperature and storage time, and for the significance of the statistical model are shown in Table 16-5.
For each food category, the five exposure models inputs are assigned a sensitivity rank based upon the magnitude of their respective F-values. For example, for deli salad, the initial LM concentration has an F-value of 8,890, which is a ratio of 2.7 greater than the input with the second highest F-value. Furthermore, the F-value for the initial LM concentration is statistically significant. Thus, the initial LM concentration is inferred to be the most sensitive input for this food category, and it is substantially more sensitive than the second ranked input of growth rate at 5 $^{0}$C. The F-value of the second ranked input is a ratio of 1.7 greater than that for the third ranked input. Thus, the second ranked variable is more sensitive than the third ranked input of storage temperature. The F-value of the third ranked input is a ratio of 4.5 greater than that for the 4$^{th}$ ranked input. Thus, the third ranked variable is substantially more sensitive than the fourth ranked input of serving size. The F-value of the fourth ranked input is a ratio of 2 greater than that for the fifth ranked input of storage time. The smallest F-value of 187 for storage time is statistically significant. The F-value of the interaction term for storage time and storage temperature of 198 is statistically significant and is more than F-value of fifth ranked input of storage time. However, the F-value for the interaction term is still a ratio of 1.7 smaller than the fourth ranked input and it is not as important in this case compared to the top ranked individual inputs.

In comparing results for different food categories, it is apparent that the initial LM concentration is the top ranked input in all cases and that the F-values for initial LM concentration are substantially larger by a ratio of 2 to 10 compared to the second largest inputs. The second most important input varies from one food category to another. For milk, pate, and smoked seafood, the storage temperature is the second most important variable. The F-value for storage temperature in these three cases is substantially less than that for the initial LM concentration and is substantially larger than that for the third ranked input. Thus, the relative importance of storage temperature for these three foods is unambiguous. For deli salad and fresh soft cheese, growth rate at 5 $^{0}$C is identified as the second most important input. Its F-value is

Table 16-4. Factor Levels for Co-mingled Variability and Uncertainty for the *Listeria monocytogenes* Exposure Model Inputs for Deli Salad, Fresh Soft Cheese, Milk, Pâtés, and Smoked Seafood

| Factor | Level 1[(a)] | Level 2[(b)] | Level 3[(c)] |
|---|---|---|---|
| **Deli Salad** | | | |
| Serving Size (g) | 0 to 100 | 100 to 220 | 220 to 1415 |
| Initial LM Concentration (log cfu/g) | -3.5 to -2 | -2 to -1 | -1 to 3.7 |
| Storage temperature ($^0$C) | 0 to 4 | 4 to 5 | 5 to 22 |
| Storage time (days) | 0 to 3 | 3 to 4 | 4 to 12.5 |
| Growth at 5 $^0$C (log cfu/day) | -0.5 to -0.25 | -0.25 to 0 | 0 to 0.2 |
| **Fresh Soft Cheese** | | | |
| Serving Size (g) | 0 to 25 | 25 to 30 | 30 to 250 |
| Initial LM Concentration (log cfu/g) | -2.7 to -1.5 | -1.5 to 0 | 0 to 9 |
| Storage temperature ($^0$C) | 0 to 4 | 4 to 5 | 5 to 22 |
| Storage time (days) | 0 to 2.5 | 2.5 to 5.5 | 5.5 to 25 |
| Growth at 5 $^0$C (log cfu/day) | -0.16 to -0.06 | -0.06 to 0.22 | 0.22 to 0.33 |
| **Milk Levels** | | | |
| Serving Size (g) | 0 to 240 | 240 to 360 | 360 to 3910 |
| Initial LM Concentration (log cfu/g) | -4 to -2 | -2 to -0.5 | -0.5 to 7.5 |
| Storage temperature ($^0$C) | 0 to 4 | 4 to 5 | 5 to 22 |
| Storage time (days) | 0 to 2.75 | 2.75 to 6 | 6 to 15.5 |
| Growth at 5 $^0$C (log cfu/day) | 0.02 to 0.175 | 0.175 to 0.35 | 0.35 to 0.8 |
| **Pâtés Levels** | | | |
| Serving Size (g) | 0 to 55 | 55 to 100 | 100 to 455 |
| Initial LM Concentration (log cfu/g) | -3 to -1.4 | -1.4 to 1.6 | 1.6 to 9.1 |
| Storage temperature ($^0$C) | 0 to 4 | 4 to 5 | 5 to 22 |
| Storage time (days) | 0 to 5 | 5 to 10 | 10 to 40 |
| Growth at 5 $^0$C (log cfu/day) | -032 to 0.16 | 0.16 to 0.3 | 0.3 to 0.75 |
| **Smoked Seafood Levels** | | | |
| Serving Size (g) | 0 to 30 | 30 to 80 | 80 to 145 |
| Initial LM Concentration (log cfu/g) | -2.5 to -1 | -1 to 2.5 | 2.5 to 9 |
| Storage temperature ($^0$C) | 0 to 4 | 4 to 5 | 5 to 22 |
| Storage time (days) | 0 to 3 | 3 to 7 | 7 to 28 |
| Growth at 5 $^0$C (log cfu/day) | 0.025 to 0.08 | 0.08 to 0.18 | 0.18 to 0.41 |

(a) Lower 0.2 percentile range; (b) 0.2 to 0.8 percentile range; (c) 0.8 to 1.0 percentile range

Table 16-5. ANOVA Results for Co-mingled Variability and Uncertainty for the Main Effects of Individual Inputs, Interactions, and Statistical Model Significance for the *Listeria monocytogenes* Exposure Model for Deli Salad, Fresh Soft Cheese, Milk, Pâtés, and Smoked Seafood

| Factor | F Value | Pr > F | Significant | Rank |
|---|---|---|---|---|
| **Deli Salad** | | | | |
| Serving Size (g) | 367 | <.0001 | Yes | 4 |
| Initial LM Concentration (log cfu/g) | 8886 | <.0001 | Yes | 1 |
| Storage temperature ($^0$C) | 1885 | <.0001 | Yes | 3 |
| Storage time (days) | 187 | <.0001 | Yes | 5 |
| Growth at 5 $^0$C (log cfu/day) | 3301 | <.0001 | Yes | 2 |
| Interaction (temp*time) | 198 | <.0001 | Yes | |
| Model | 1124 | <.0001 | Yes | |
| **Fresh Soft Cheese** | | | | |
| Serving Size (g) | 651 | <.0001 | Yes | 4 |
| Initial LM Concentration (log cfu/g) | 13768 | <.0001 | Yes | 1 |
| Storage temperature ($^0$C) | 834 | <.0001 | Yes | 3 |
| Storage time (days) | 624 | <.0001 | Yes | 5 |
| Growth at 5 $^0$C (log cfu/day) | 2053 | <.0001 | Yes | 2 |
| Interaction (temp*time) | 126 | <.0001 | Yes | |
| Model | 1717 | <.0001 | Yes | |
| **Milk** | | | | |
| Serving Size (g) | 436 | <.0001 | Yes | 5 |
| Initial LM Concentration (log cfu/g) | 9009 | <.0001 | Yes | 1 |
| Storage temperature ($^0$C) | 3134 | <.0001 | Yes | 2 |
| Storage time (days) | 1210 | <.0001 | Yes | 3 |
| Growth at 5 $^0$C (log cfu/day) | 710 | <.0001 | Yes | 4 |
| Interaction (temp*time) | 318 | <.0001 | Yes | |
| Model | 1182 | <.0001 | Yes | |
| **Pâtés** | | | | |
| Serving Size (g) | 109 | <.0001 | Yes | 5 |
| Initial LM Concentration (log cfu/g) | 10145 | <.0001 | Yes | 1 |
| Storage temperature ($^0$C) | 5217 | <.0001 | Yes | 2 |
| Storage time (days) | 1089 | <.0001 | Yes | 4 |
| Growth at 5 $^0$C (log cfu/day) | 1182 | <.0001 | Yes | 3 |
| Interaction (temp*time) | 274 | <.0001 | Yes | |
| Model | 1682 | <.0001 | Yes | |

(Continued on next page.)

Table 16-5. Continued

| Factor | F Value | Pr > F | Significant | Rank |
|---|---|---|---|---|
| **Smoked Seafood** | | | | |
| Serving Size (g) | 262 | <.0001 | Yes | 5 |
| Initial LM Concentration (log cfu/g) | 22536 | <.0001 | Yes | 1 |
| Storage temperature ($^0$C) | 2090 | <.0001 | Yes | 2 |
| Storage time (days) | 413 | <.0001 | Yes | 3 |
| Growth at 5 $^0$C (log cfu/day) | 410 | <.0001 | Yes | 4 |
| Interaction (temp*time) | 187 | <.0001 | Yes | |
| Model | 2054 | <.0001 | Yes | |

substantially larger than that for the third ranked input of storage temperature. Thus the F-value of all variables are in different ranges and the rankings are robust.

In many cases, it is possible to clearly identify the least important inputs and to do so without any ambiguity. The clearest example is for serving size for pâtés, which has an F-value of 109 that is a factor 10 smaller than highest ranked input of initial LM concentration. In this case, the F-value is statistically significant but is small compared to the F-value of the highest ranked input. As another example, for deli salad, the fifth ranked input of storage time has an F-value of 187 compared to an F-value of 8,890 for the highest ranked input. Thus, although there is a statistically significant relationship between the mean responses for dose with respect to different levels of storage time, the mean response for dose is substantially more strongly influenced by other inputs.

Storage time and serving size are the two of the least important variables. Storage time is the least important variable for deli salad and fresh soft cheese whereas serving size is the least important input variable for milk, pâtés and smoked seafood. Furthermore, the F-values for the fifth ranked inputs range from 109 for pâtés to 624 for fresh soft cheese, compared to F-values ranging from 8,890 to 22,500 for the top ranked inputs in each food category. Thus, on a comparative basis, it is possible to clearly distinguish the importance of the top ranked inputs vis-à-vis inputs with substantially smaller F-values.

The F-values for the interaction between storage temperature and storage time are typically low, ranging from 126 for fresh soft cheese to 318 for milk, when compared to the F-values for the top ranked inputs. These results suggest that the interaction term is not of

substantial importance, even though it is more important than storage time alone and is statistically significant.

The conclusions based upon the information presented in Tables 16-4 and 16-5 are summarized here:

- For all five food groups, the initial LM concentration is by far the most important input.
- For deli salad, the growth rate at 5 $^0$C, storage temperature, and serving size are of decreasing importance, respectively. The interaction between storage time and storage temperature is relatively unimportant. High storage temperatures are conducive to higher doses regardless of the storage time.
- For fresh soft cheese, growth at 5 $^0$C is clearly the second most important variable. Storage time, serving size and storage temperature are relatively not as important.
- For milk, storage temperature is clearly the second most important input. All variables have well separated F-values and hence the rankings are unambiguous. Storage time, growth rate at 5 $^0$C and serving size are of decreasing importance.
- For pâtés and smoked seafood, storage temperature is clearly the second most important input. Storage time and growth rate at 5 $^0$C are substantially less important than storage temperature. The interaction term between storage time and storage temperature was of relatively small importance.

Although the interaction between storage time and storage temperature appears to be of less importance than the individual effect of many of the inputs, a more detailed evaluation of the interaction between these two inputs may reveal combinations of levels for these two factors that lead to substantial changes in the mean response. The ANOVA results for contrasts of these two factors are given in Table 16-6 for all five selected food categories. The contrasts are constructed in manner similar to Section 16.1.1.

The results demonstrate that the change in the mean response when storage time changes from a high to a low level is much larger when storage temperature is at a high level than when storage temperature is at a medium or low level. For example, for fresh soft cheese the estimate for high storage temperature level and contrast between high and low storage time is 0.818, which is much larger than the estimate for the same contrast in storage time for medium or low levels of storage temperature. Furthermore, the estimate for all three possible contrasts for storage time associated with a high level of storage temperature is higher than the estimates for

Table 16-6. ANOVA Results for Variability and Uncertainty for Contrasts of Storage Time and Storage Temperature for the *Listeria monocytogenes* Exposure Model for Deli Salad, Fresh Soft Cheese, Milk, Pâtés, and Smoked Seafood

| TEMPERATURE LEVELS [(a)] | CHANGE IN TIME LEVELS | | |
|---|---|---|---|
| | **High to Low** | **High to Medium** | **Medium to Low** |
| **Deli Salad** | | | |
| Low | 17.23, Yes, -0.065 | 0.17, No, -0.007 | 11.45, Yes, -0.058 |
| Medium | 3.75, No, -0.040 | 0.01, No, -0.002 | 3.52, No, -0.037 |
| High | 1059.67, Yes, -0.584 | 581.11, Yes, -0.457 | 53.92, Yes , -0.127 |
| **Fresh Soft Cheese** | | | |
| Low | 73.42, Yes, 0.204 | 127.37, Yes, 0.209 | 0.02, No, -0.004 |
| Medium | 575.63, Yes, 0.676 | 441.65, Yes, 0.530 | 34.63, Yes, 0.146 |
| High | 1103.04, Yes, 0.818 | 691.06, Yes, 0.611 | 114.41, Yes, 0.206 |
| **Milk** | | | |
| Low | 133.99, Yes, 0.350 | 163.88, Yes, 0.302 | 3.65, No, 0.048 |
| Medium | 785.79, Yes, 1.120 | 550.86, Yes, 0.794 | 121.99, Yes, 0.326 |
| High | 2818.94, Yes, 1.807 | 1146.29, Yes, 1.016 | 1153.57 , Yes, 0.791 |
| **Pâtés** | | | |
| Low | 157.32, Yes, 0.389 | 171.12 , Yes, 0.300 | 8.73, Yes, 0.089 |
| Medium | 756.09, Yes, 1.179 | 210.81, Yes, 0.471 | 361.42, Yes, 0.708 |
| High | 2406.69 , Yes, 1.801 | 998.73, Yes, 0.925 | 875.53, Yes, 0.875 |
| **Smoked Seafood** | | | |
| Low | 63.58, Yes, 0.209 | 13.68, Yes, 0.076 | 27.15, Yes, 0.132 |
| Medium | 165.60, Yes, 0.481 | 150.65, Yes, 0.354 | 13.90, Yes, 0.127 |
| High | 1236.51, Yes, 1.205 | 442.95, Yes, 0.648 | 479.54, Yes, 0.557 |

(a) For Low, Medium and High levels of temperature the F value, statistical significance and estimate of difference between mean responses for each contrast are shown.

medium or low levels of storage temperature. This result is biologically plausible since higher temperatures are associated with higher growth rates of *Listeria monocytogenes*. If the temperature is in the low or medium levels, the growth rate is slow enough that there is less sensitivity to storage time.

The conclusions regarding each temperature level corresponding to the three contrasts, for each of the five food categories are summarized here:

For deli salad,

- At low temperatures, it is important to constrain the storage time to a low level otherwise there is a significant change in response.
- At medium temperatures, controlling the storage time does not yield any significant change in mean response.
- At high temperatures, if storage time changes from one level to another the response will change significantly.

For fresh soft cheese,

- At low and medium temperature, exposure is sensitive to changes between medium and high storage time and between low and high storage time. However, a change from low to medium storage time is not significant.
- At medium and high temperatures, if storage time changes from one level to another the response will change significantly.

For milk,

- At low temperatures, differences between either low or medium storage time versus high storage time are important, but differences between low and medium storage times are not.
- At medium and high temperatures, any control measure that changes the storage time from one level to another will result in significant change in response.

For pâtés,

- At all temperature levels, any control measure that changes the storage time at retail from one level to another will yield significant change in response.

For smoked seafood,

- At all temperature levels, any control measure that changes the storage time at retail from one level to another will yield significant change in response.

The top two inputs for all food groups except fresh soft cheese were the same for both the variability only and co-mingled variability and uncertainty dataset cases. Also, the least important input in each of the five food categories was same for both datasets. Hence the difference in the distributions assumed in the two probabilistic analyses did not substantially affect the ranking of the most and least important inputs.

## 16.2 Application of ANOVA to the Dose Response Module

ANOVA was applied to the dose response module for five selected food groups, including deli salad, fresh soft cheese, milk, pâtés, and smoked seafood. For each food group, the inputs of interest include uncertainty in dose adjustment factor, exposure period, virulence susceptibility uncertainty, mouse lethality uncertainty, and population exposure. The output of interest is mortality in the neonatal sub-population. Thus for ANOVA application, the factors in the dose response module are the five inputs and the response variable is the output.

The dose response module incorporates only uncertainty. This section focuses on application of ANOVA based upon uncertainty in inputs. The point estimates of some input variable such as virulence susceptibility, mouse lethality and fraction of population exposed are probability values and not actual value. No interactions among variables were considered to be of any practical significance. The input assumptions regarding factor levels for the inputs for the five food categories are given in Table 16-7. The results of ANOVA for each individual factor and for the significance of the statistical model are shown in Table 16-8.

For each food category, the five exposure models inputs are assigned a sensitivity rank based upon the magnitude of their respective F-values. For example, for deli salad, the dose adjustment factor has an F-value of 5.3, which is significant. All other inputs are insignificant. For four of five food categories, the dose adjustment factor is ranked first and is statistically significant. For fresh soft cheese and smoked seafood the fraction of population exposed is also a significant input. The dose adjustment factor is ranked second to the fraction of population exposed in the smoked seafood category. F-values for most of the inputs for any given food category were statistically significant.

## 16.3 Summary of the Result of Application of ANOVA

The results for the exposure module included cases based upon variability only as well as both variability and uncertainty co-mingled in a single dimension. The latter typically has a wider range of variation for a given input than does the former. The results from both cases were

Table 16-7. Factor Levels for Uncertainty Only for the *Listeria monocytogenes* Dose-Response Model Inputs for Deli Salad, Fresh Soft Cheese, Milk, Pâtés, and Smoked Seafood

| Factor | Level 1[(a)] | Level 2[(b)] | Level 3[(c)] |
|---|---|---|---|
| **Deli Salad** | | | |
| Dose Adjustment Factor | 5 to 7 | 7 to 8 | 8 to 11.5 |
| Exposure Period | 1 to 7 | 7 to 18 | 18 to 30 |
| Virulence Susceptibility | 0 to 0.2 | 0.2 to 0.8 | 0.8 to 1 |
| Mouse Lethality | 0 to 0.2 | 0.2 to 0.8 | 0.8 to 1 |
| Fraction of Population Exposed | 0 to 0.2 | 0.2 to 0.8 | 0.8 to 1 |
| **Fresh Soft Cheese** | | | |
| Dose Adjustment Factor | 5 to 6.5 | 6.5 to 8 | 8 to 11.5 |
| Exposure Period | 1 to 7.5 | 7.5 to 18 | 18 to 30 |
| Virulence Susceptibility | 0 to 0.2 | 0.2 to 0.8 | 0.8 to 1 |
| Mouse Lethality | 0 to 0.2 | 0.2 to 0.8 | 0.8 to 1 |
| Fraction of Population Exposed | 0 to 0.2 | 0.2 to 0.8 | 0.8 to 1 |
| **Milk** | | | |
| Dose Adjustment Factor | 5 to 7 | 7 to 8 | 8 to 11.5 |
| Exposure Period | 1 to 7 | 7 to 18 | 18 to 30 |
| Virulence Susceptibility | 0 to 0.2 | 0.2 to 0.8 | 0.8 to 1 |
| Mouse Lethality | 0 to 0.2 | 0.2 to 0.8 | 0.8 to 1 |
| Fraction of Population Exposed | 0 to 0.2 | 0.2 to 0.8 | 0.8 to 1 |
| **Pâtés** | | | |
| Dose Adjustment Factor | 5 to 7 | 7 to 8 | 8 to 11.5 |
| Exposure Period | 1 to 6 | 6 to 18 | 18 to 30 |
| Virulence Susceptibility | 0 to 0.2 | 0.2 to 0.8 | 0.8 to 1 |
| Mouse Lethality | 0 to 0.2 | 0.2 to 0.8 | 0.8 to 1 |
| Fraction of Population Exposed | 0 to 0.2 | 0.2 to 0.8 | 0.8 to 1 |
| **Smoked Seafood** | | | |
| Dose Adjustment Factor | 5 to 7 | 7 to 8 | 8 to 11.5 |
| Exposure Period | 1.25 to 7 | 7 to 18 | 18 to 30 |
| Virulence Susceptibility | 0 to 0.2 | 0.2 to 0.8 | 0.8 to 1 |
| Mouse Lethality | 0 to 0.2 | 0.2 to 0.8 | 0.8 to 1 |
| Fraction of Population Exposed | 0 to 0.2 | 0.2 to 0.8 | 0.8 to 1 |

(a) Lower 0.2 percentile range
(b) 0.2 to 0.8 percentile range
(c) 0.8 to 1.0 percentile range

Table 16-8. ANOVA Results for Uncertainty Only for the Main Effects of Individual Inputs, Interactions, and Statistical Model Significance for the *Listeria monocytogenes* Dose-Response Model for Deli Salad, Fresh Soft Cheese, Milk, Pâtés, and Smoked Seafood

| Factor | F Value | Pr > F | Significant | Rank |
|---|---|---|---|---|
| **Deli Salad** | | | | |
| Dose Adjustment Factor | 5.3 | 0.0052 | Yes | 1 |
| Exposure Period | 1.6 | 0.1951 | No | - |
| Virulence Susceptibility | 0.7 | 0.5093 | No | - |
| Mouse Lethality | 0.1 | 0.8955 | No | - |
| Fraction of Population Exposed | 2.9 | 0.0577 | No | - |
| Model | 1.6 | 0.0049 | Yes | |
| **Fresh Soft Cheese** | | | | |
| Dose Adjustment Factor | 20.4 | <.0001 | Yes | 1 |
| Exposure Period | 0.2 | 0.8028 | No | - |
| Virulence Susceptibility | 1.0 | 0.3580 | No | - |
| Mouse Lethality | 1.8 | 0.1705 | No | - |
| Fraction of Population Exposed | 4.3 | 0.0131 | Yes | 2 |
| Model | 2.1 | <.0001 | Yes | |
| **Milk** | | | | |
| Dose Adjustment Factor | 13.8 | <.0001 | Yes | 1 |
| Exposure Period | 3.8 | 0.0230 | Yes | 2 |
| Virulence Susceptibility | 0.9 | 0.4172 | No | - |
| Mouse Lethality | 0.5 | 0.6403 | No | - |
| Fraction of Population Exposed | 2.5 | 0.0817 | No | - |
| Model | 2.1 | <.0001 | Yes | |
| **Pâtés** | | | | |
| Dose Adjustment Factor | 17.4 | <.0001 | Yes | 1 |
| Exposure Period | 1.1 | 0.3401 | No | - |
| Virulence Susceptibility | 0.5 | 0.6059 | No | - |
| Mouse Lethality | 0.9 | 0.3987 | No | - |
| Fraction of Population Exposed | 1.9 | 0.1471 | No | - |
| Model | 3.0 | <.0001 | Yes | |
| **Smoked Seafood** | | | | |
| Dose Adjustment Factor | 4.9 | 0.0075 | Yes | 2 |
| Exposure Period | 0.3 | 0.7243 | No | - |
| Virulence Susceptibility | 1.8 | 0.1591 | No | - |
| Mouse Lethality | 0.1 | 0.9469 | No | - |
| Fraction of Population Exposed | 6.8 | 0.0011 | Yes | 1 |
| Model | 3.0 | <.0001 | Yes | |

typically similar, implying that the identification of sensitive inputs was robust with respect to differences in the distributions assumed for each input. However, there were some cases, such as for fresh soft cheese, in which there were noticeable differences in ranking based upon the two different probabilistic simulation approaches. In principle, ANOVA could be applied as part of a two dimensional simulation, as demonstrated for the *E. coli* model in Chapter 5. In order to apply ANOVA as part of a two dimensional probabilistic simulation, it is necessary to consistently define factor levels for each of the uncertainty iterations.

ANOVA is a generalized statistical method for sensitivity analysis. ANOVA was applied to both the exposure and dose-response modules of the *Listeria monocytogenes* model for each of five food categories. The F-value of each individual factor was used to rank order the importance of the factors in the case of statistically significant F-values. For F-values that were not statistically significant, a judgment was made that the model output was insensitive to the respective input. The F-values of interaction terms were evaluated to determine whether interactions were statistically significant. In many cases, interactions were evaluated in more detail based upon the use of contrasts. The analysis of the contrasts typically revealed that variation in one input may lead to a significantly higher mean value of the response conditional on values of another input.

The case studies results illustrate that it is possible to distinguish sensitive inputs from insensitive (or statistically insignificant) ones. Furthermore, it is possible to discriminate importance among the statistically significant inputs. For example, it is often possible to identify an input that is clearly substantially more important than another input, even though both may have a statistically significant influence on the output.

ANOVA is able to deal with models that are nonlinear or that contain thresholds. The ability to analyze contrasts, for example, illustrates the capability of ANOVA to identify how the output is sensitive to a particular input conditional on a specific level of another input. The division of the domain of an input into levels provides a method for evaluating how the output may respond differently to different ranges of values of the input. Differences in responses for different levels could be because of nonlinearity or the presence of thresholds.

# 17 APPLICATION OF CLASSIFICATION REGRESSION TREES (CART) TO THE *LISTERIA MONOCYTOGENES* MODEL

The purpose of this chapter is to apply CART to the *Listeria monocytogenes* model. CART is discussed in Section 2.2.5. The *Listeria monocytogenes* model is discussed in Chapter 12. CART does not assume any specific model and it groups data into a number of classes. The data within a class are more homogeneous in response than the total dataset. A key consideration in this chapter is to evaluate the within food category rankings of the inputs in different classes of exposure values, especially at high exposure values. This chapter is divided into three major sections. The first section focuses on the application of CART to the exposure module of the *Listeria monocytogenes* model. The second section focuses on the application of CART to the dose-response module of the *Listeria monocytogenes* model. The third section summarizes the main findings from the application of CART to the *Listeria monocytogenes* model.

In order to apply CART to different modules of the *Listeria monocytogenes* food safety risk assessment model, S-PLUS© Version 6.1 is used. This software has the ability to perform CART analysis on a dataset, using a Graphical User Interface (GUI). Moreover, both qualitative and quantitative inputs can be addressed using the specific options of the software. The software provides option for specifying branch lengths as proportional to the deviance that the partition accounted for or a constant length. Due to the difficulty reading the graph for the deviance proportional case, the constant length option was used. In addition to visualization of the regression tree, an additional method for ranking sensitive inputs was evaluated based upon estimation of the reduction in total deviance that can be attributed to each input selected for inclusion in the regression tree.

## 17.1 Application of CART to the Exposure Module

CART was applied to the exposure module for five selected food groups, including deli salad, fresh soft cheese, milk, pâtés, and smoked seafood. For each food group, the inputs of interest include serving size in grams, initial LM concentration in log cfu/g, storage temperature in $^0$C, storage time in days, and growth potential at 5 $^0$C in log cfu/day. The output of interest is the dose value corresponding to each meal serving size simulated.

The following subsection focuses on application of CART based upon only variability in inputs and the corresponding estimated variability in dose. Section 17.1.2 focuses upon

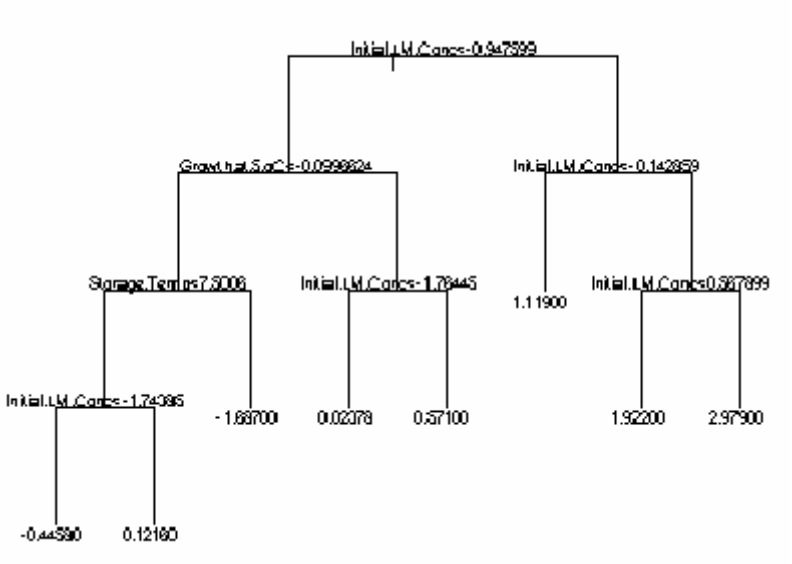



Figure 17-1. Regression Tree For Variability Only Inputs For *Listeria monocytogenes* Exposure Module For Deli Salad.

application of CART when both variability and uncertainty are considered together, but not distinguished, to represent a randomly selected individual. The details of dataset generation are described in Section 12.9. The purpose of conducting both analyses is to gain insight into how the contribution of uncertainty changes, if at all, to the results of sensitivity analysis compared to a case in which only variability is considered. The regression trees for each of the five food categories are presented in the results.

CART classifies the data according to the mean exposure levels of homogeneous subgroups. To gain further insight about the rankings of inputs within different exposure classes, regression analysis can be applied to the datasets corresponding to classes created by CART. However, for demonstration purposes, regression analysis is performed only on classes of the dataset obtained for deli salad under the variability only condition. Similar analysis can be performed on any other regression tree.

### 17.1.1 Sensitivity Analysis of Exposure Module Based Upon Variability Only

This section covers application of CART to the exposure module when only variability was considered. The uncertainty dimension of inputs was fixed to point values as described in Section 12.9. CART was performed for each of five selected food categories. The results of CART for deli salad, fresh soft cheese, milk, pâtés and smoked seafood are presented in Figures 17-1 to 17-5, respectively.

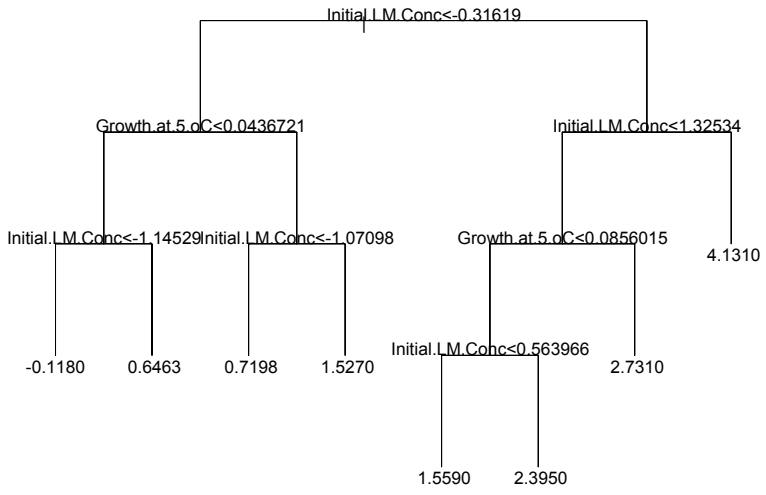



Figure 17-2. Regression Tree For Variability Only Inputs For *Listeria monocytogenes* Exposure Module For Fresh Soft Cheese.

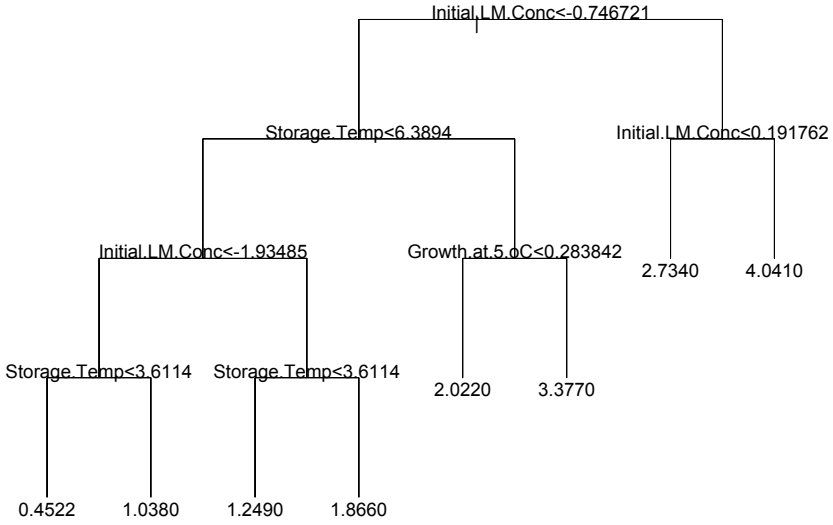



Figure 17-3. Regression Tree For Variability Only Inputs For *Listeria monocytogenes* Exposure Module For Milk.

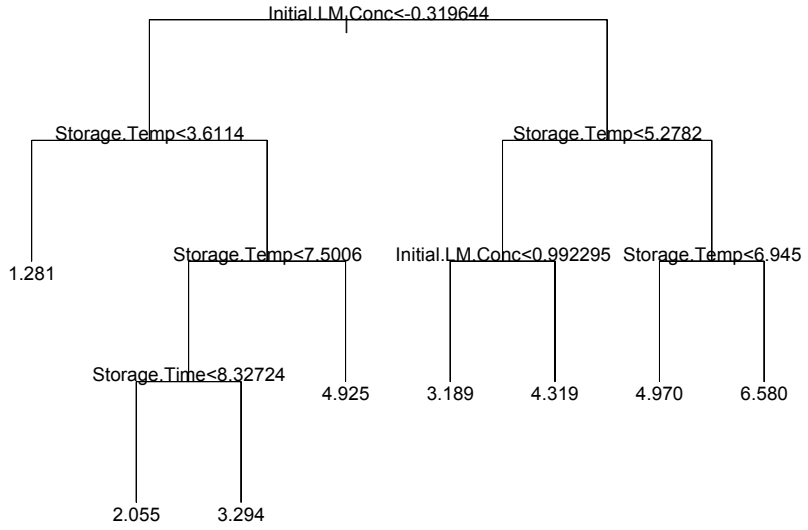



Figure 17-4. Regression Tree For Variability Only Inputs For *Listeria monocytogenes* Exposure Module For Pâtés.

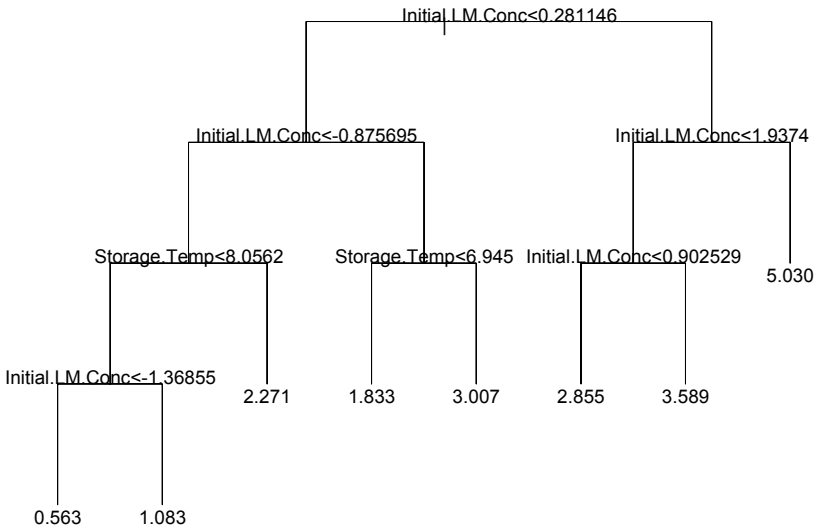



Figure 17-5. Regression Tree For Variability Only Inputs For *Listeria monocytogenes* Exposure Module For Smoked Seafood.

The results of CART are presented in form of regression trees. The mean response for each of the resulting classes is specified at the end of each leaf node. The tree interpretation is explained using the regression tree of Figure 17-1 for deli salad as an example. The initial LM concentration is the root node and it splits the dataset into two. The data with initial LM concentration less than -0.95 are the part of the left branch and the rest of the data are part of the right branch. At the next level, the right branch is further split at an initial LM concentration of -0.14 and the left branch is split based upon growth at 5 $^0$C of -0.10. The data were split into eight subgroups as indicated at the terminal leaves of the tree.

The path through the tree from the root to a terminal leaf gives a classification rule for a subgroup from the root to a leaf node. For example, the classification rule for data in the highest mean exposure subgroup for deli salad is: initial LM concentration $\geq$ 0.57. Similarly, the classification rule for lowest mean exposure subgroup is: initial LM concentration < -1.7, growth at 5 $^0$C < -0.10 and storage temperature < 7.5.

The within food category rank for some inputs can be inferred from the regression tree. For each food category, the exposure model inputs were assigned a rank based upon their position in the regression tree. However, the regression tree can give insight about relative ranking only in specific cases. For example, the input at root node is usually identified as the most important input. This input is selected at the root node because it is associated with the single largest incremental reduction in deviance. For all five food categories the initial LM concentration is the root node and hence is assumed to be the most important input. When the left and right branches of the root node are further split based upon the same input at both branches, then the input used for the second split may be ranked second. For example, for pâtés, the second most important input is inferred to be storage temperature as indicated in Figure 17-4. However, when the left and right branches use different inputs for the next split, a clear ranking may not be apparent. In such cases, the variable selected that discriminates among the highest exposures might be considered more important from a risk management perspective.

As an alternative sensitivity index, the amount of contribution of each input to the reduction of the total deviance is considered. An example for such an approach is presented for deli salad. The dataset obtained for deli salad has a total deviance of 3977. If no limit is imposed on the number of nodes, the regression tree can capture 76 percent of the total deviance. Table 17-1 summarizes the amount of contribution of each input to the reduction of total deviance.

Table 17-1. Reduction in Deviance Associated with Selected Inputs in the Regression Tree Generated in the Exposure Module of Deli Salad for the Variability Only Analysis

| Level of the Tree | Selected Inputs in the Regression Tree | | | |
|---|---|---|---|---|
| | **Serving Size** | **Initial LM Concentration** | **Storage Temperature** | **Growth at 5 $^0$C** |
| **1st Level** | | 1547 | | |
| **2nd Level** | | 267 | | 283 |
| **3rd Level** | 53 | 144<br>74 | 233 | |
| **4th Level** | 58 | 96 | | 35<br>35 |
| **5th Level** | 36<br>10 | 13<br>27 | 41 | |
| **6th Level** | 20 | | 15 | 13<br>9 |
| **Sum** | 177 | 2168 | 289 | 375 |
| **Percent of Contribution (1)** | 4.5 | 54.5 | 7.3 | 9.4 |
| **Rank** | 4 | 1 | 3 | 2 |

(3) Total deviance of the dataset is 3977. The amount of deviance captured by the regression tree is 3010.

Four inputs were selected in the regression tree. These inputs include serving size, initial LM concentration, storage temperature and growth at 5 $^0$C. Table 17-1 indicates that there were 6 levels in the regression tree. An input may appear several times under different branches of a given level. Each such appearance is denoted with a numerical entry in this table. The initial LM concentration was selected in the first split of the tree and led to the largest incremental reduction in deviance. The division of the dataset based on the first split reduced the total deviance approximately 40 percent. At the second level of the tree, two inputs were selected as the basis for branching, and the reductions in deviance attributable to each of these two inputs are comparable to each other but are substantially smaller than the reduction in deviance for the first level split. At each level except the first two, three different inputs were selected for the splits.

Initial LM concentration, storage temperature, growth at 5 $^0$C and serving size were selected seven, three, six and five times, respectively. However, serving size was selected mostly in the lower levels. For each input in Table 17-1 the percent of contribution to the total reduction in the deviance is identified. These contributions vary between 5.5 and 54.5 percent. The selected inputs in the regression tree are ranked based on their contribution to the reduction in total deviance reduction. The ranking indicates that the initial LM concentration reduces the total deviance by 54.5% and is easily the most important input. Growth at 5 $^0$C, storage temperature

and serving size follow as the second, third and fourth important inputs with 9.4%, 7.3% and 4.5% contribution to reduction in total deviance. Based on the rankings the input selected at the first splitting node has the highest contribution to reduction in the total deviance. The number of times an input appears in the tree does not correspond to the degree of its relative rank. However, the level with which an input first appears in the tree has an appropriate association with its importance in this case.

Based on the input contribution to the reduction of the total deviance, the inputs can be divided into three groups: (1) initial LM concentration is the most important input; (2) the growth at 5 $^0$C, storage temperature, and serving size are of approximately comparable but minor importance; and (3) storage time is not important since it was not selected for inclusion in the tree.

As shown in Figure 17-1, the exposure values were divided into eight subgroups using CART. To further analyze these datasets, regression analysis was applied to three classes with the highest mean exposure levels for the example of deli salad. Regression analysis was also applied on the remaining five subgroups taken together. The results are given in Table 17-2. The results indicate that the importance of inputs varies with the subgroup. For example, initial LM concentration was ranked first for the highest mean exposure subgroup, second for the second highest mean exposure subgroup and fourth for the third highest mean exposure subgroup. The low ranking for the latter subgroup is because the initial LM concentration input in this subgroup is bounded between -0.95 and -0.14 and hence the scope of variability for the initial LM concentration is restricted. Other inputs, such as growth at 5 $^0$C and serving size, are not bounded and this likely contributes to their higher ranking for the third subgroup. However storage time was identified as the least important input, irrespective of the subgroup. The magnitude of the parameter estimates for inputs within a subgroup were often similar and in some cases the 95 percent confidence intervals overlapped. For example, for the three highest mean exposure subgroups, the 95 percent confidence intervals overlap for serving size and growth at 5 $^0$C. Thus these two inputs were of comparable importance.

The top two important inputs for the highest mean exposure level of 2.98 were initial LM concentration and serving size. However, the top two-ranked inputs are different for lower exposure subgroups. For example the top two inputs for third highest exposure dataset

Table 17-2. Parameter Estimates and Within Food Category Rankings for the *Listeria monocytogenes* Exposure Model Inputs for Data Subgroups With Mean Exposure Values of 2.98, 1.92, 1.12 and for Data Subgroups Consisting of the Rest Data for Deli Salad

| Variable Name | Parameter Estimate | 95% Confidence Interval | Rank Within The Food Category | Statistically Significant |
|---|---|---|---|---|
| **Deli Salad Dataset Subgroup With Mean Exposure Level of 2.98** | | | | |
| Serving Size in g | 0.38 | 0.29 to 0.48 | 2 | Yes |
| Initial LM Concentration in log cfu/g | 0.67 | 0.58 to 0.76 | 1 | Yes |
| Storage temperature in $^0$C | -0.15 | -0.24 to -0.06 | 4 | Yes |
| Storage time in days | -0.02 | -0.11 to 0.08 | 5 | No |
| Growth at 5 $^0$C in log cfu/day | 0.30 | 0.21 to 0.40 | 3 | Yes |
| **Deli Salad Dataset Subgroup With Mean Exposure Level of 1.92** | | | | |
| Serving Size in g | 0.55 | 0.47 to 0.63 | 1 | Yes |
| Initial LM Concentration in log cfu/g | 0.29 | 0.21 to 0.38 | 3 | Yes |
| Storage temperature in $^0$C | -0.28 | -0.37 to -0.19 | 4 | Yes |
| Storage time in days | -0.14 | -0.23 to -0.06 | 5 | Yes |
| Growth at 5 $^0$C in log cfu/day | 0.54 | 0.45 to 0.62 | 2 | Yes |
| **Deli Salad Dataset Subgroup With Mean Exposure Level of 1.12** | | | | |
| Serving Size in g | 0.43 | 0.38 to 0.48 | 2 | Yes |
| Initial LM Concentration in log cfu/g | 0.26 | 0.21 to 0.32 | 4 | Yes |
| Storage temperature in $^0$C | -0.37 | -0.43 to -0.32 | 3 | Yes |
| Storage time in days | -0.17 | -0.22 to -0.12 | 5 | Yes |
| Growth at 5 $^0$C in log cfu/day | 0.48 | 0.43 to 0.53 | 1 | Yes |
| **Deli Salad Dataset Subgroup Consisting of Remaining Dataset** | | | | |
| Serving Size in g | 0.34 | 0.32 to 0.37 | 4 | Yes |
| Initial LM Concentration in log cfu/g | 0.55 | 0.53 to 0.58 | 1 | Yes |
| Storage temperature in $^0$C | -0.35 | -0.38 to -0.33 | 3 | Yes |
| Storage time in days | -0.13 | -0.16 to -0.11 | 5 | Yes |
| Growth at 5 $^0$C in log cfu/day | 0.49 | 0.47 to 0.51 | 2 | Yes |

corresponding to a mean exposure of 1.12 are growth at 5 $^0$C and serving size. The top inputs for the high exposure level are generally of the most important concern to risk managers. Thus, the initial LM concentration is suggested as the most important input because it was the top ranked input for the highest mean exposure subgroup.

For fresh soft cheese the initial LM concentration is the most important input. Growth at 5 $^0$C is the only other input appearing in the regression tree. Storage times, serving size and growth at 5 $^0$C are of less importance compared to initial LM concentration and growth at 5 $^0$C because they were not selected in the tree. The classification rule for highest exposure subgroup is solely determined by initial LM concentration and it corresponds to initial LM concentration $\geq$ 1.33.

The regression tree for milk included the initial LM concentration, storage temperature and growth at 5 $^0$C as the important variables. Storage time and serving did not appear in the tree and hence were of lesser importance. The classification rule for the highest exposure subgroup is given by initial LM concentration $\geq$ 0.19.

In the case of pâtés, the initial LM concentration was the most important input. Storage temperature can be considered as second most important variable as it participates in both the left and right branches for the second split. The classification rule for highest exposure subgroup is given by an initial LM concentration $\geq$ -0.32 and storage temperature $\geq$ 6.95 $^0$C.

The initial LM concentration is the most important input for the smoked seafood category as it participates in five of the seven splits, including the root node. Storage temperature is the only other input participating in the tree. The serving size, storage time and growth at 5 $^0$C are of less importance than both the initial LM concentration and storage temperature because these three inputs are not included in the tree. The classification rule for the highest mean exposure subgroup is given by an initial LM concentration $\geq$ 1.94.

The comparisons of ranks obtained corresponding to the three subgroups and full the data is shown in Table 17-3. Clearly, the rankings vary with the subgroups corresponding to different exposure estimate levels.

In comparing results for different food categories, it is apparent that the initial LM concentration is the top ranked input in four of the five cases. For pâtés and smoked seafood, storage temperature is the second most important input. The inputs participating in the tree are considered to be more important than the ones absent. Thus the initial LM concentrations,

Table 17-3. Comparative Rankings of Different Subgroups Created Using CART and Full Dataset For Exposure Module of Deli Salad Food Category in *Listeria monocytogenes* Model

| Variable Name | Exposure Dataset Subgroups | | | | |
|---|---|---|---|---|---|
| | Highest Mean Exposure | Second Highest Mean Exposure | Third Highest Mean Exposure | Remaining Classes Taken Together | Full |
| Serving Size in g | 2 | 1 | 2 | 4 | 3 |
| Initial LM Concentration in log cfu/g | 1 | 3 | 4 | 1 | 1 |
| Storage temperature in $^0$C | 4 | 4 | 3 | 3 | 4 |
| Storage time in days | 5 | 5 | 5 | 5 | 5 |
| Growth at 5 $^0$C in log cfu/day | 3 | 2 | 1 | 2 | 2 |

growth at 5 $^0$C and storage temperature are the three most important source of variability for deli salad. For fresh soft cheese, the initial LM concentration and growth at 5 $^0$C are more important than other inputs. For milk, the initial LM concentration, storage temperature and growth at 5 $^0$C are more important than others. For pâtés, initial LM concentration, storage temperature and storage time are more important than the rest. For smoked seafood, initial LM concentration and storage temperature are more important than the others. Thus, CART separated the important inputs from unimportant inputs.

The initial LM concentration and storage temperature are important across food groups with the only exception that storage temperature is not among the top two inputs for fresh soft cheese. Growth at 5 $^0$C is important for both fresh soft cheese and milk as it appears in the two respective trees. Initial LM concentration is the most important input as it appears at the root node in all cases. Serving size is absent from all the five food categories trees and thus is an unimportant input across all food categories.

### 17.1.2 Sensitivity Analysis of Exposure Module Based Upon Both Variability and Uncertainty Co-Mingled

In this section, CART is applied to the exposure module when both variability and uncertainty are considered together in one-dimensional simulation. The details of dataset generation are described in Section 12.9. CART was performed for each of five selected food categories.

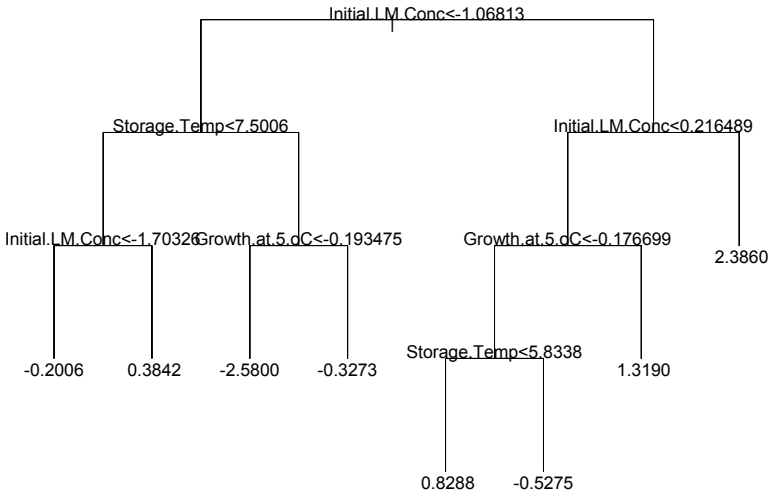
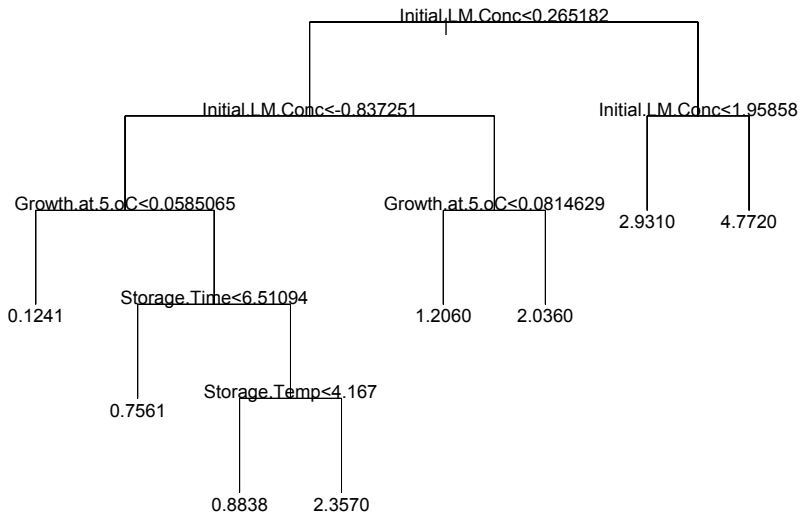



Figure 17-6. Regression Tree for Both Variability and Uncertainty in Inputs For *Listeria monocytogenes* Exposure Module For Deli Salad.

The results of CART are shown in the form of regression trees. The results of CART for each food category are shown in Figure 17-6 to 17-10 for deli salad, fresh soft cheese, milk, pâtés and smoked seafood, respectively. The mean response for each of the subgroups is specified at the end of each leaf node. For example, for deli salad, as shown in Figure 17-6, the initial LM concentration is the root node. It splits the dataset into two. The datasets with initial LM concentration less than -1.07 are part of left branch and the rest are part of the right branch. At the next level of node, the right branch is further split by an initial LM concentration of 0.22 and a storage temperature of 7.5 $^0$C splits the left branch. The initial LM concentration, storage temperature, and growth at 5 $^0$C were the only input selected as the basis for splitting the data. The dataset was repeatedly split using the cost complexity pruning option in S-PLUS© until eight subgroups were obtained. The path gives a classification rule for a subgroup from root to a leaf node. For example, the classification rules for data in the highest mean exposure subgroup, which has a mean of 2.39 is that, the initial LM concentration $\geq$ 0.22.

For deli salad, the initial LM concentration was the top ranked input. Storage temperature and growth at 5 $^0$C are the other two inputs that appear in the tree. However, it is not clear which of the two is more important with respect to each other. The inputs initial storage time and




Figures 17-7. Regression Tree Both Variability and Uncertainty in Inputs For *Listeria monocytogenes* Exposure Module For Fresh Soft Cheese.

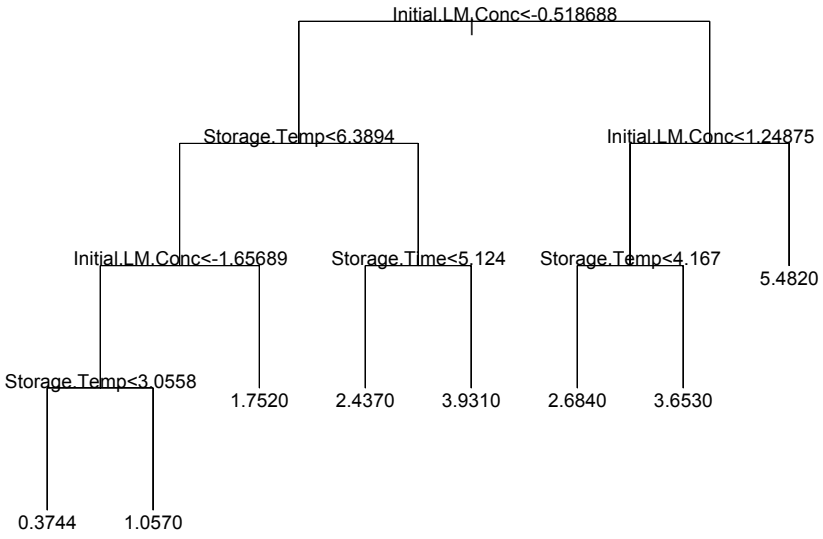



Figure 17-8. Regression Tree For Both Variability and Uncertainty in Inputs For *Listeria monocytogenes* Exposure Module For Milk.

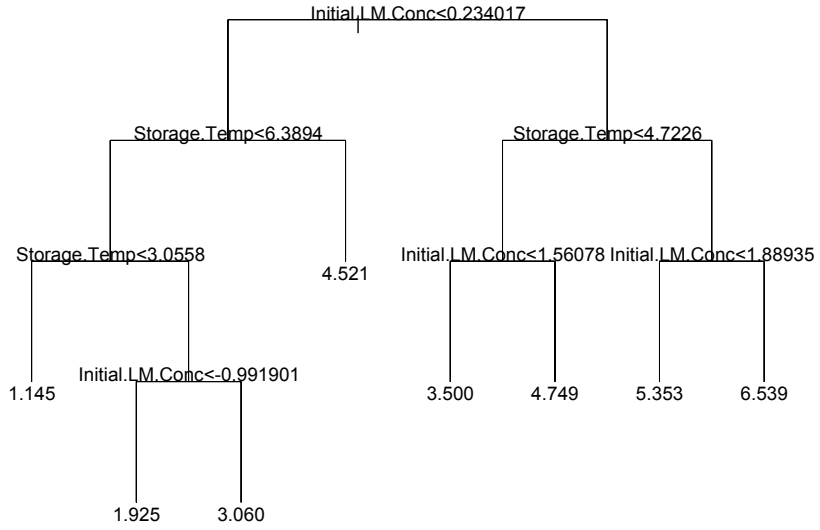



Figure 17-9. Regression Tree For Both Variability and Uncertainty in Inputs For *Listeria monocytogenes* Exposure Module For Pâtés.

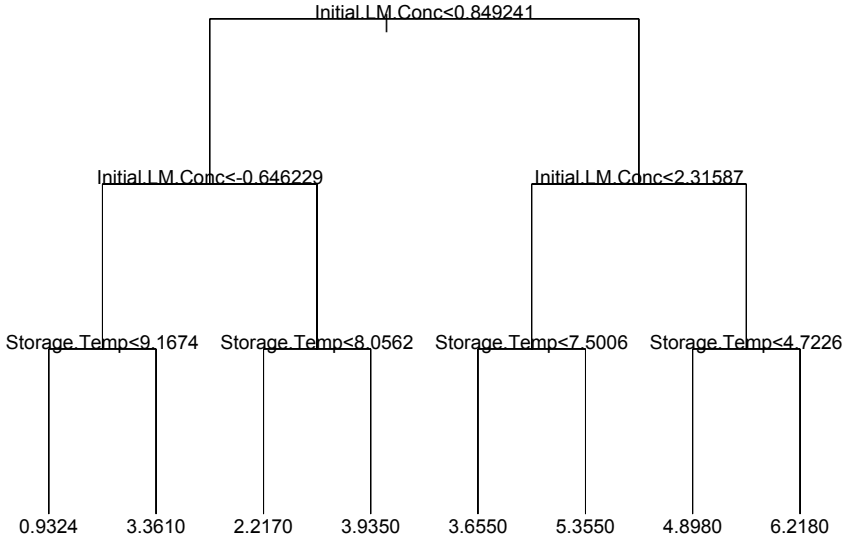



Figure 17-10. Regression Tree For Both Variability and Uncertainty in Inputs For *Listeria monocytogenes* Exposure Module For Smoked Seafood.

serving size do not appear in the tree, indicating that their contribution to overall variance in output is small or negligible compared to the three inputs that are present.

For fresh soft cheese the initial LM concentration was at the root node and hence is considered to be the most important input. The second most important input was growth at 5 $^0$C. Storage time and temperature appear in a few of the lower nodes of the tree and hence are less important compared to initial LM concentration and growth at 5 $^0$C. Serving size did not appear in the tree and therefore is deemed to be unimportant. The classification rule for highest mean exposure subgroup is that the initial LM concentration $\geq$ 1.96.

In the case of milk, the regression tree had initial LM concentration at the root node and thus this is considered to be the most important input. Storage temperature and storage time also appear in the tree. Serving size and growth at 5 $^0$C were not included in the tree and therefore are less important than initial LM concentration, storage temperature and storage time. The classification rule for highest mean exposure subgroup is that the initial LM concentration $\geq$ 1.25.

For pâtés, the initial LM concentration is the most important variable. Storage is the second ranked input. Serving sizes, growth at 5 $^0$C and storage time do not appear in the tree and hence are of less important the initial LM concentration and storage temperature.

The regression tree for smoked seafood shows clear relative importance between inputs. The initial LM concentration is the most important input. In fact it is the only input that appears in the first and second level splits. Storage temperature is the next important variable. Serving size, storage time and growth at 5 $^0$C were not included in the tree and thus are of less importance than the two present in the regression tree.

For all five food categories, the initial LM concentration was the root node and hence it was the most important input across all food categories. When the left and right branches of the root node had the same input used for next split, then the input used for second split is ranked second. For example, in the case of pâtés, the input that split the data after the first split on both the left and right branches was storage temperature. Thus storage temperature was judged to be more important than all variables except initial LM concentration. Similarly, the second most important input for fresh soft cheese and smoked seafood can be identified as growth at 5 $^0$C and storage temperature, respectively. However, when the left and right branches used different inputs for the next split, the rankings were not easily identified. Thus for deli salad and milk it is

difficult to determine which input is clearly second ranked. However, the inputs that are present in the tree are more important than those absent. Thus the initial LM concentration, growth at 5 $^0$C and storage temperature are the three most important sources of variability for deli salad. For fresh soft cheese, serving size is of least importance, as it does not appear in the tree. For milk, initial LM concentration, storage temperature and time are more important than others. For pâtés, initial LM concentration and storage temperature are more important than the rest. For smoked seafood, initial LM concentration and storage temperature are more important than the others. Thus, CART separates the important inputs from less important inputs.

## 17.2 Application of CART to the Dose Response Module

CART was applied to the dose response module for five selected food groups, including deli salad, fresh soft cheese, milk, pâtés, and smoked seafood. For each food group, the inputs of interest include uncertainty in dose adjustment factor, exposure period, virulence susceptibility uncertainty, and mouse lethality uncertainty and population exposure. The output of interest is mortality ascertained to the food group in the neonatal sub-population. The dose response module incorporates only uncertainty. This section focuses on application of CART based upon uncertainty in inputs. The dataset used for application of CART to the dose response module was described in Section 12.9. The results of CART are shown in form of regression trees. The mean response for each of the classes is specified at the end of each leaf node. The results of CART for deli salad, fresh soft cheese, milk, pâtés and smoked seafood are presented in Figures 17-11 to 17-15, respectively.

For deli salad and smoked seafood the fraction of population exposed was the most important input. For milk and pâtés, dose adjustment is the most important input. Clearly, dose adjustment and fraction of population exposed were more important than other inputs not appearing in the tree. The mouse lethality input was present for the deli salad and pâtés regression trees but was absent from the regression trees for other food categories.

For pâtés, the dose adjustment input was the root node. The dataset was split into two based on dose adjustment. Those sets of inputs with dose adjustment less than 7.12 were part of left branch and the rest of the right branch. At the next level, the right branch was further split with respect to dose adjustment factor of 7.25 and left branch for exposure period of 14.4. The dataset was repeatedly split until eight classes were formed. The classification rule for data in

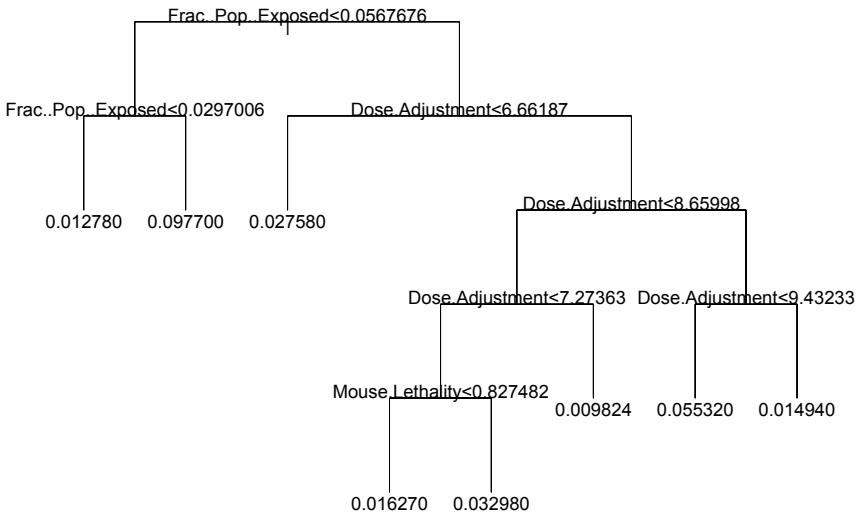



Figure 17-11. Regression Tree For Uncertainty Only Inputs For *Listeria monocytogenes* Dose Response Module For Deli Salad.

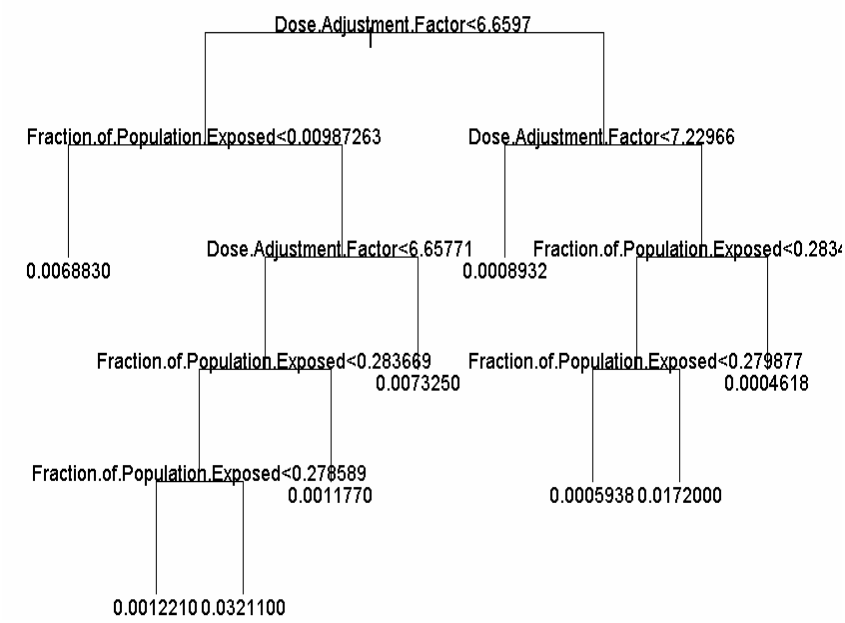



Figure 17-12. Regression Tree For Uncertainty Only Inputs For *Listeria monocytogenes* Dose Response Module For Fresh Soft Cheese.

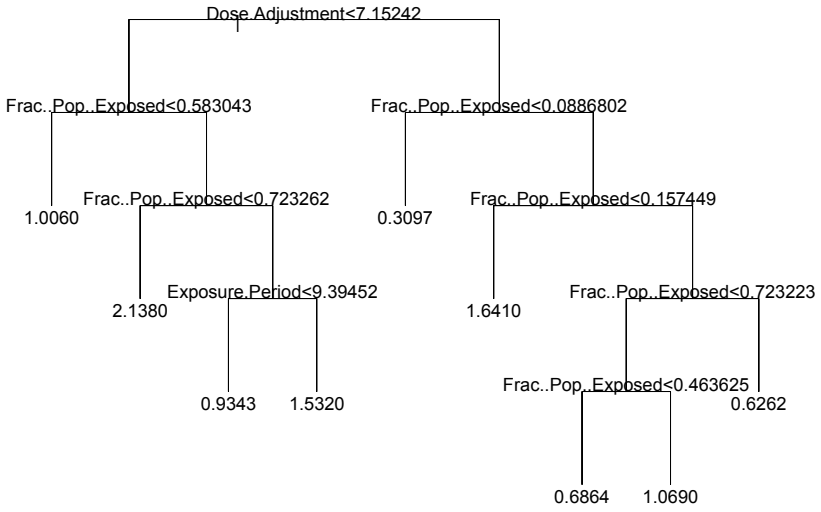



Figure 17-13. Regression Tree For Uncertainty Only Inputs For *Listeria monocytogenes* Dose Response Module For Milk.

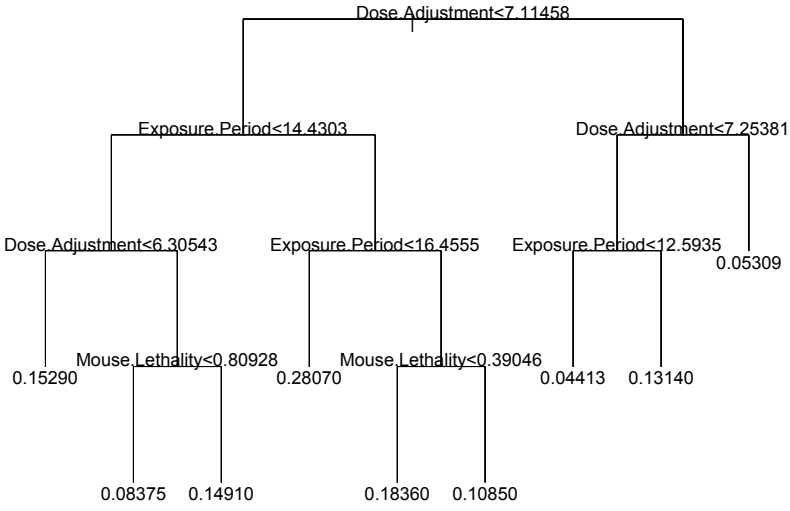



Figure 17-14. Regression Tree For Uncertainty Only Inputs For *Listeria monocytogenes* Dose Response Module For Pâtés.

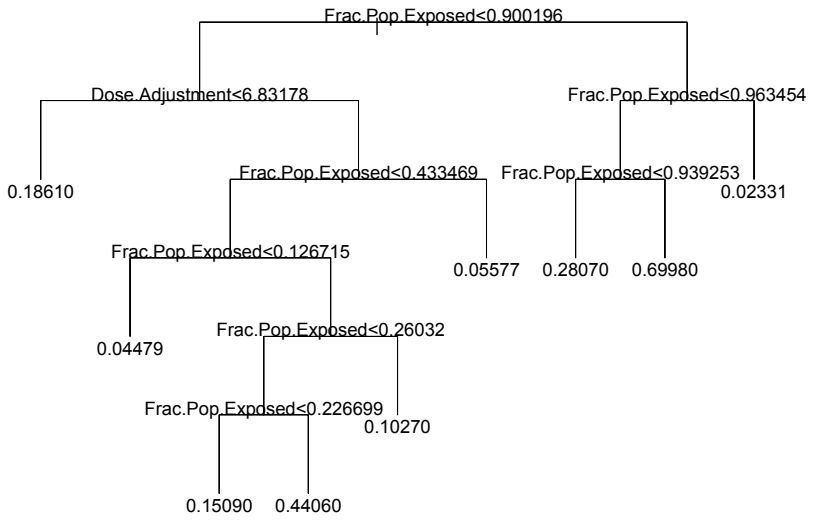



Figure 17-15. Regression Tree For Uncertainty Only Inputs For *Listeria monocytogenes* Dose Response Module For Smoked Seafood.

highest mean mortality class is given by dose adjustment < 7.12, exposure period ≥ 16.5 and mouse lethality < 0.39.

The dose adjustment factor and fraction of population exposed consistently appeared as top inputs the regression tree of each of the five food categories. Hence these two inputs are deemed more important than others not appearing in the tree.

## 17.3 Summary of the Results of Application of CART

In this chapter CART was applied to specific modules of the *Listeria monocytogenes* model in order to identify the most important factors influencing the response of selected outputs. CART does not assume specific functional relation between the model inputs and the model response. Hence, for models that have nonlinearity or thresholds application of CART does not force any under estimation or overestimation regarding the sensitivity of the output to each input.

CART does not present any specific sensitivity index. The ranking of the inputs in CART is based on visualization of the regression tree and judgment. In some cases, the application of other sensitivity analysis methods as a complement to CART is needed to gain insight regarding

the rank of each input. Because CART does not produce sensitivity index similar to those of methods such as ANOVA or regression analysis, it is difficult to automate CART for application to many iterations. For example, when variability is simulated separately for multiple realizations of uncertainty in a two-dimensional probabilistic framework, it is difficult to summarize and compare the results of the CART analysis for each of the 100 iterations. Thus, the lack of a quantitative sensitivity index is a practical limitation that makes it difficult to automate CART for use with two-dimensional probabilistic analysis.

Although there is no direct sensitivity parameter, the sum of the deviance reduction by each input in the formation of the tree may be used as a sensitivity index. Such an estimate can allow for comparison of inputs present in the tree. However, this index is not a direct output from S-PLUS GUI interface. Manual calculation of the indices for comparison is tedious. Future development of an automated procedure for calculation of reduction in deviance is recommended.

CART identified the most important inputs in all the cases. It separated important inputs from less important variables. Initial LM concentration in the exposure module always appeared as the top most important input. For both variability only and, variability and uncertainty co-mingled cases in the exposure module, the input serving size never appeared in the regression tree. Thus, serving size was less important compared to other variables. CART also helped classify data according to output levels. Since CART did not assume any model, the results were independent of linear/non-linear response assumptions.

The regression trees for dose response module separated dose adjustment factor and fraction of population exposed as the top inputs across all food categories. Although mouse lethality and exposure period appeared in the regression trees for deli salad and pâtés, respectively, they were always low in the tree. Thus, they were not as important as the dose adjustment factor or fraction of population exposed.

The classification rules obtained using CART analyses were useful in grouping data into various exposure subgroups. The variables that participated in these classification rules are expected to be more important inputs than the ones not participating. In most cases, the initial LM concentration was the only factor that decided the highest exposure class. For example, in the case of the deli salad food group, when only variability was considered an initial LM concentrate greater than 0.57 solely decided if the response is in highest exposure class.

# 18 APPLICATION OF SCATTER PLOTS TO THE *LISTERIA MONOCYTOGENES* MODEL

The purpose of this chapter is to apply scatter plots to the *Listeria monocytogenes* model. Scatter plots are discussed in Section 2.3.1. The *Listeria monocytogenes* model is discussed in Chapter 12. Scatter plots are used as a visual aid to identify patterns in the response corresponding to important inputs.

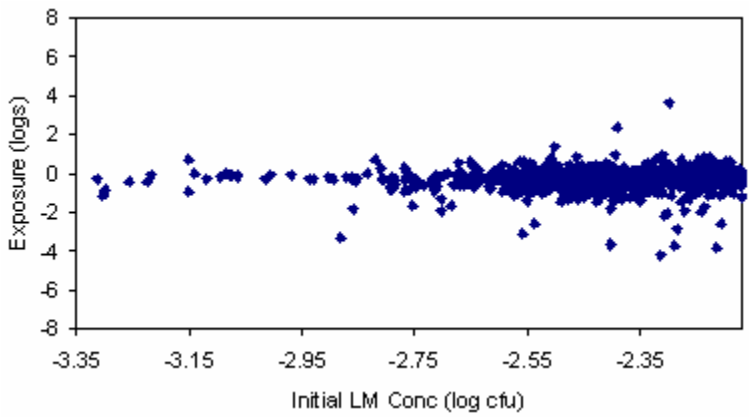
For different food categories and modules of the model, scatter plots were prepared only for those inputs that were ranked high by statistical methods presented in Chapters 15 to 17. The prioritization of effort was necessitated by the difficulties in generating scatter plots in terms of time spent and interpretation. To get a better local view of any specific pattern in response, three graphs were prepared for each variable corresponding to the lower tail, middle and upper tail region of the variable. For each exposure module input in each food category, the ranges for lower tail, middle and upper tail region are given in Table 16-1, corresponding to the levels used in ANOVA.

This chapter is divided into three major sections. The first section focuses on the application of scatter plots to the exposure module of the *Listeria monocytogenes* model. The second section focuses on the application of scatter plots to the dose-response module of the *Listeria monocytogenes* model. The third section summarizes the main findings from the application of scatter plots to the *Listeria monocytogenes* model.

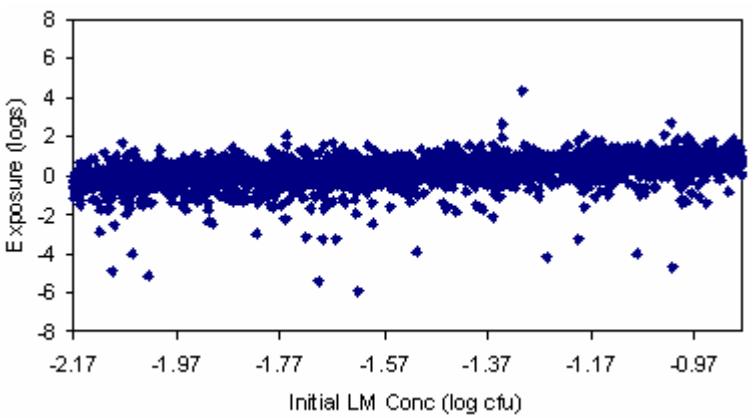
## 18.1 Application of Scatter Plots to the Exposure Module

Scatter plots were applied to the exposure module for five selected food groups, including deli salad, fresh soft cheese, milk, pâtés, and smoked seafood. For each food group, the inputs of interest include serving size in grams, initial LM concentration in log cfu/g, storage temperature in $^0$C, storage time in days, and growth potential at 5 $^0$C in log cfu/day. The output of interest was the dose value corresponding to each meal serving size simulated. Scatter plots were applied to the dataset obtained for cases where only variability was considered. The details of dataset generation are described in Section 12.9.

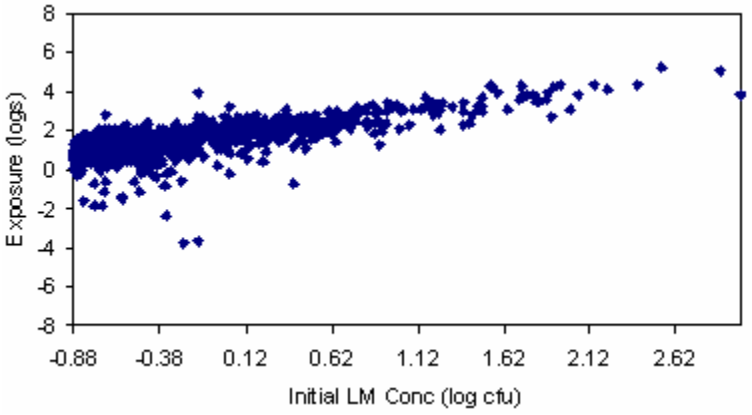
Regression Analysis, ANOVA and CART identified initial LM concentration and growth at 5 $^0$C as the first and second ranked inputs for deli salad. Thus these two inputs were selected for evaluation using scatter plots. The scatter plots for deli salad are shown in Figures 18-1 and 18-2 for initial LM concentration and growth at 5 $^0$C, respectively.




Lower Tail Region

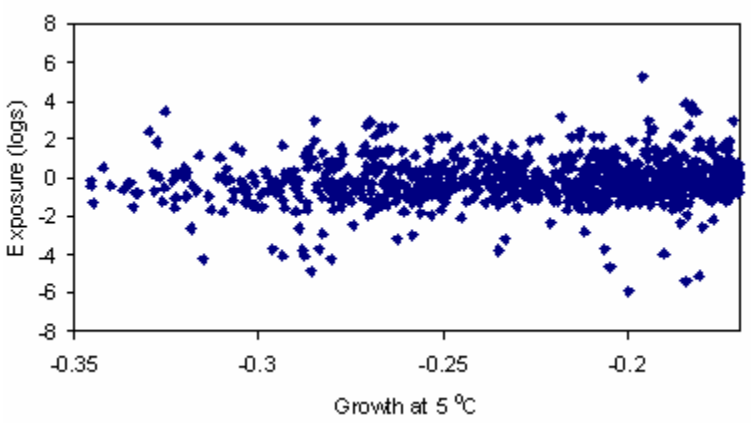



Middle Region

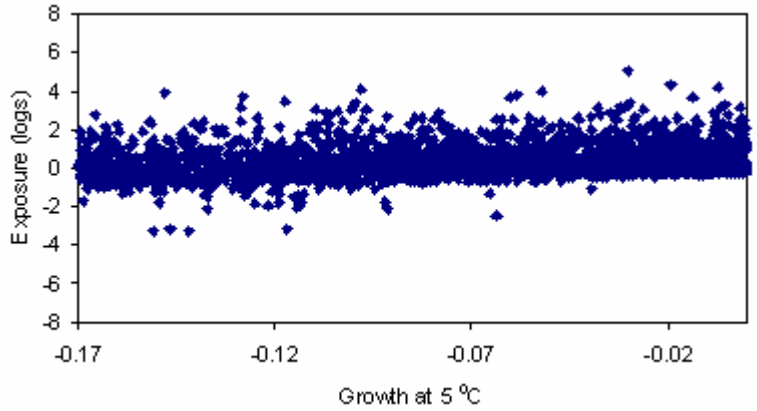



Upperr Tail Region

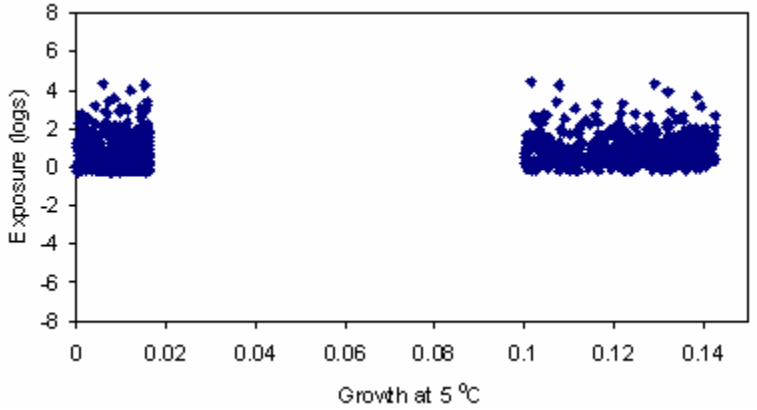
Figure 18-1. Scatter Plots of Exposure logs For Initial *Listeria monocytogenes* Concentration in Exposure Module of Deli Salad Food category.




Lower Tail Region




Middle region




Upper Tail Region

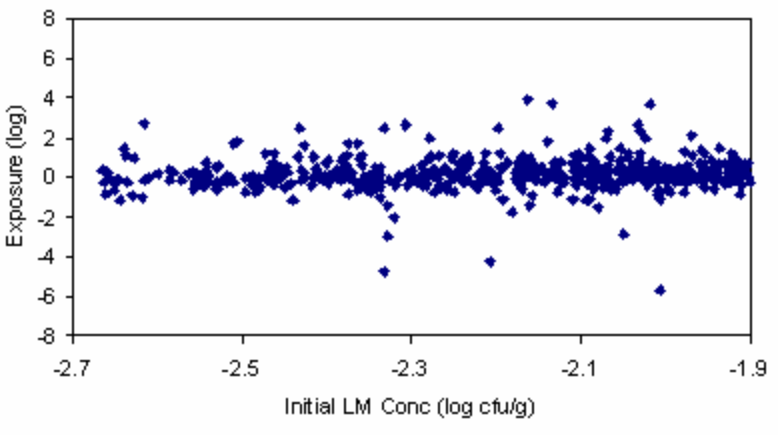
Figure 18-2. Scatter Plots of Exposure logs For Growth at 5 $^{0}$C in Exposure Module of Deli Salad Food category.

The lower tail scatter plots for initial LM concentration showed approximately constant response. The response was centered at 0 logs exposure value. Most of the values were between -2 to 2 logs, with a few outliers. The middle region scatter plots for initial LM concentration showed a linear increase in the mean response as the initial LM concentration increased. The exposure values were between -2 to 2 logs and the mean predicted exposure increased from -0.5 logs to 1 as the initial concentration increased from -2.17 to -0.97 log cfu. At the upper tail end of the initial LM concentration, the response showed an approximately linear increase. The exposure values in the tail region ranged from -4 logs to 6 logs over all values of initial LM concentration from -0.88 to 3.12 log cfu. The mean predicted exposure increased from 1 log to 6 logs. Thus an initial LM concentration below -0.88 logs, ensured that the exposure value would lie below 2 logs.

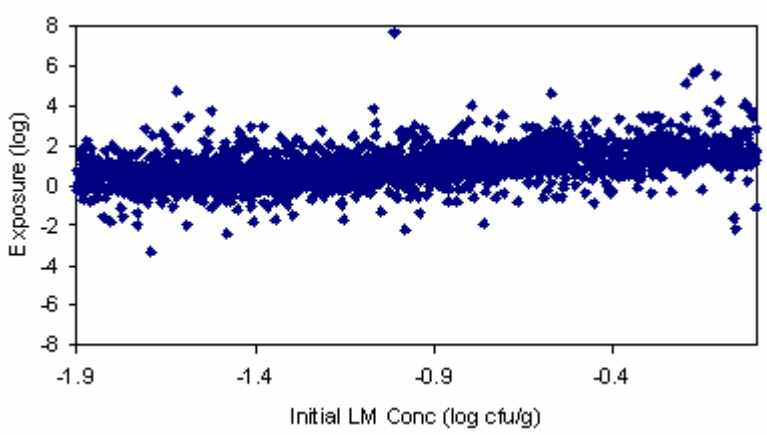
The scatter plots corresponding to growth at 5 $^0$C for deli salad showed that an increase in growth at 5 $^0$C alone does not change the exposure significantly. Also, the mean responses at most of the point estimates are same. However, in the lower tail region of the input growth at 5 $^0$C, the exposure values varied from -6 logs to 6 logs for growth at 5 $^0$C from -0.35 to -0.15 log cfu/day. Thus even when growth at 5 $^0$C was low, other inputs had a significant effect on variation in exposure. The exposure values were found to always lie above -4 logs and -0.5 logs for growth at 5 $^0$C in middle and upper tail region, respectively. There were no points in the range 0.02 logs to 0.1 logs of growth at 5 $^0$C as the distribution was modeled such.

Regression Analysis and CART identified initial LM concentration and growth at 5 $^0$C as the first and second ranked input for fresh soft cheese. The scatter plots for fresh soft cheese are shown in Figure 18-3 and 18-4 for initial LM concentration and growth at 5 $^0$C, respectively.

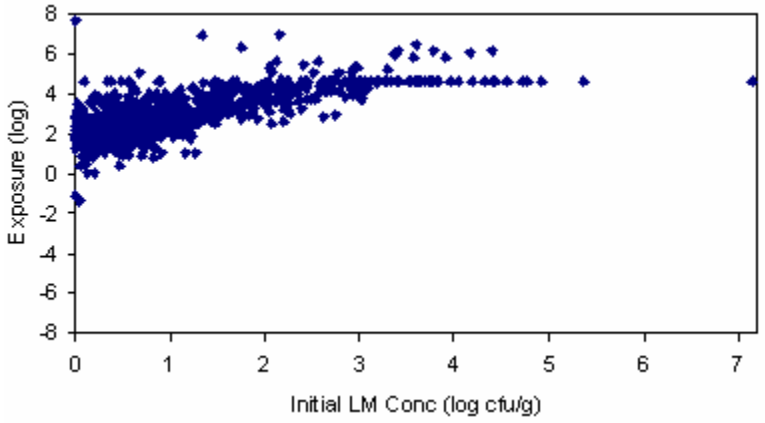
The scatter plots for lower tail region of initial LM concentration showed a response centered at approximately 0 logs exposure value. Most of the values were between -1 to 2 logs but there were a few outliers. The outliers represented the case where the growth after retail was large enough to give an exposure value of 4 logs. The scatter plots for the middle region of initial LM concentration showed an approximately linear increase in the mean response as the initial LM concentration increased. The exposure values were between -3 to 4 logs and the mean predicted exposure increased from 1 log to 2 logs as the growth at 5 $^0$C changed from -1.9 l to 0.1 log cfu/g. At the upper tail end of the initial LM concentration, the mean exposure values increased linearly from 0 to 5logs and saturated at 5 to 6 logs for initial LM concentration greater




Lower Tail Region

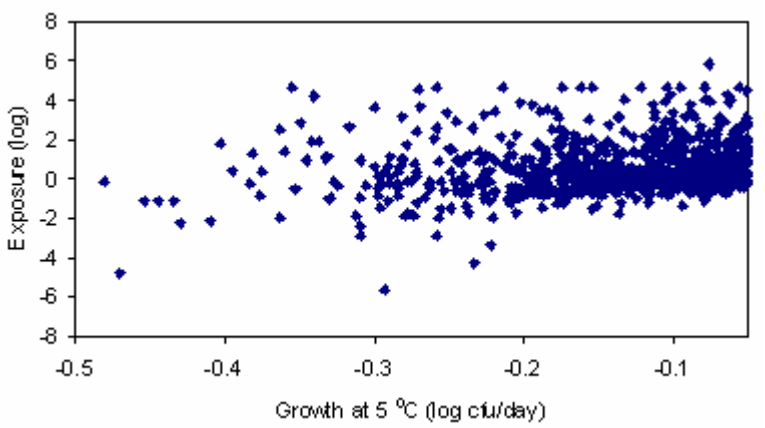



Middle Region

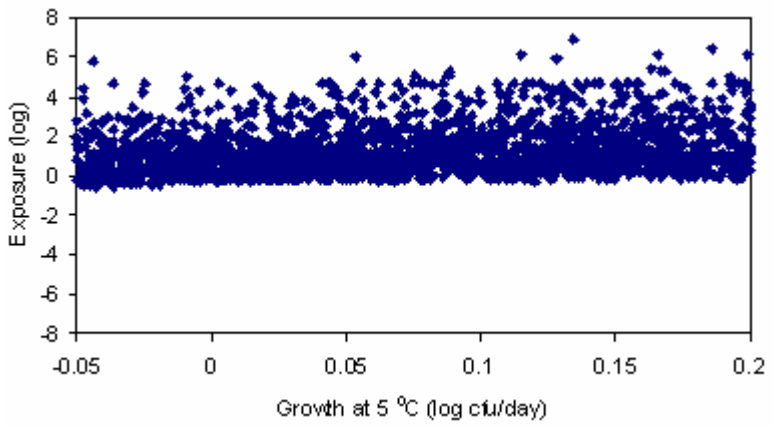



Upper Tail Region

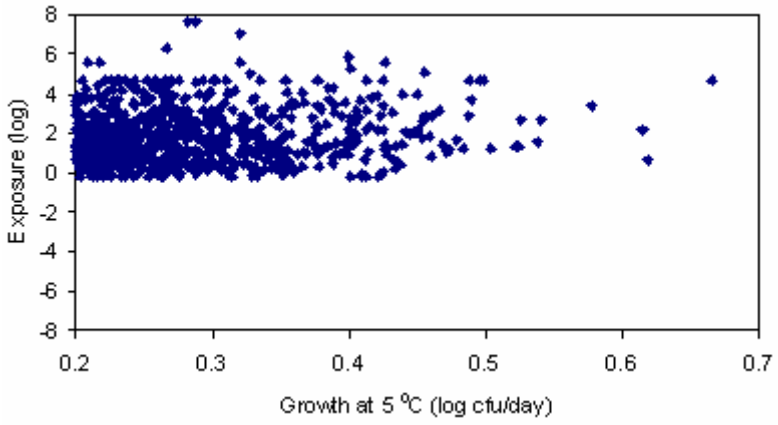
Figure 18-3. Scatter Plots of Exposure logs For Initial *Listeria monocytogenes* Concentration in Exposure Module of Fresh Soft Cheese Food category.




Lower Tail Region




Middle Region




Upper Tail Region

Figure 18-4. Scatter Plots of Exposure logs For Growth at 5 $^{0}$C in Exposure Module of Fresh soft CheeseFood category.

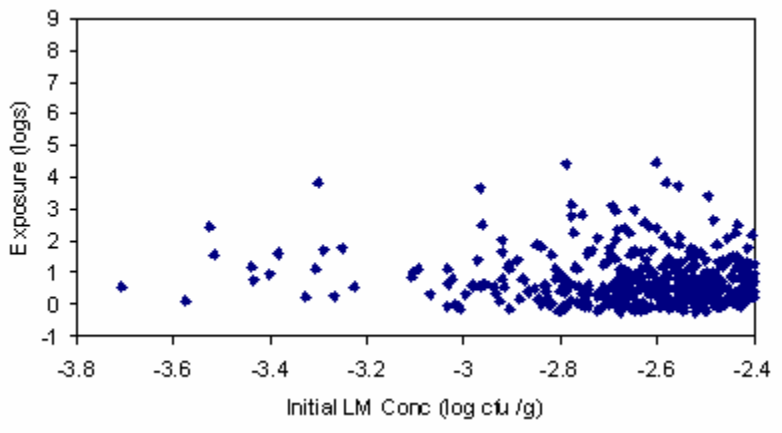
than 2.5 log cfu/g. An initial LM concentration greater than 0 logs typically had exposure values above 0 logs.

The scatter plots corresponding to growth at 5 $^0$C showed large variations in the exposure values. In the lower tail region, the exposure values ranged from -6 to 6 logs. However, when the growth at 5 $^0$C was in the middle region then the exposure values were always above -0.5 logs. In this region, exposure values ranged from -0.5 logs to 6 logs for all values of growth at 5 $^0$C. The mean exposure was approximately 1 log and did not change substantially over the middle region. For upper tail region the mean response was typically 2 logs. Thus, although growth at 5 $^0$C ensured a corresponding lower bound on exposure values, the upper bound was decided by other factors and hence the large variations were observed.

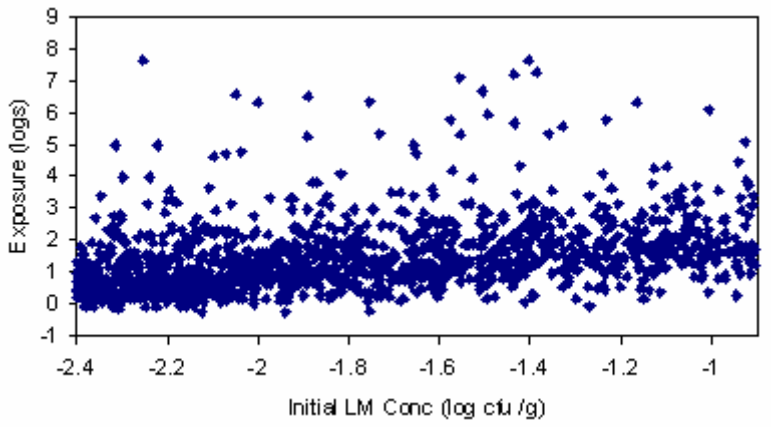
Regression Analysis, ANOVA and CART identified initial LM concentration and storage temperature as the first and second ranked input for milk. The scatter plots for milk are shown in Figure 18-5 and 18-6 for inputs initial LM concentration and storage temperature, respectively.

The lower tail scatter plots for initial LM concentration showed a response centered at 0.5 logs exposure value. Most of the values were between -0.5 to 2 logs with a few outliers. The outliers indicated that the growth after retail was significant enough to give a high exposure level of 5 logs. The middle region scatter plots for initial LM concentration showed a linear increase in the mean response as the initial LM concentration increased from -2.4 to -0.8 log cfu/g. The exposure values were mostly between -0.5 to 8 logs but the mean predicted exposure changed from 0.5 logs to 1 logs. The variations in exposure estimates for the middle region were high. Thus, the medium level initial concentration combined with good post retail growth conditions to result in high exposure value of 8 logs. At the tail end of the initial LM concentration the exposure values increased linearly up to 7 to 8 logs. The mean predicted exposure at point estimates of initial LM concentration ranged from 2 to 6 logs. Thus an initial LM concentration above -0.9 logs almost always ensured an exposure value above 0.5 logs.

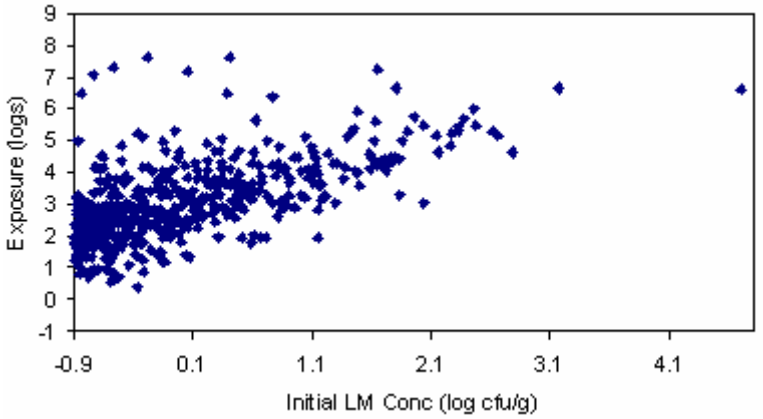
The scatter plots corresponding to storage temperature showed large variations in the exposure values. The plots reflect that the distribution for storage temperature was stepped and not continuous. In the lower tail region the exposure values ranged from -0.5 logs to 7 logs. The middle region response was similar to low region. However, when the storage temperature was in the upper tail region, the mean exposure values increased from 2 logs to 4 logs. It can be concluded that high storage temperature result in high exposure values for milk.




Lower Tail Region

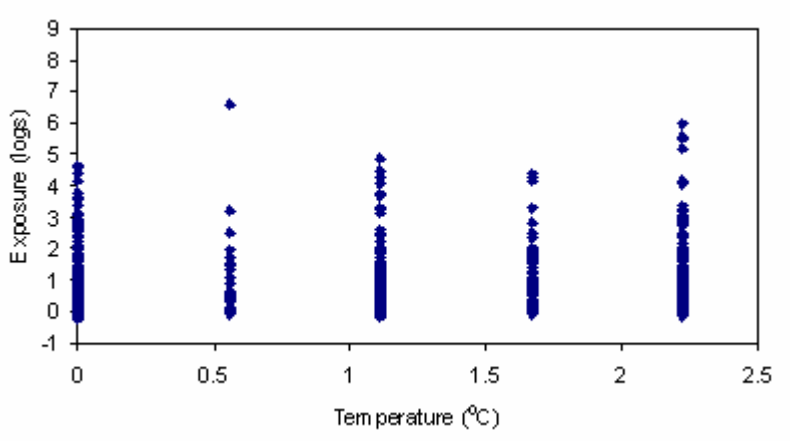



Middle Region

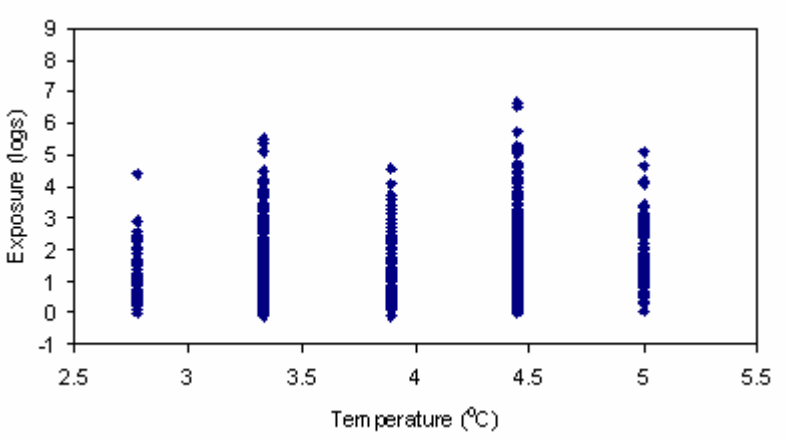



Upper Tail Region

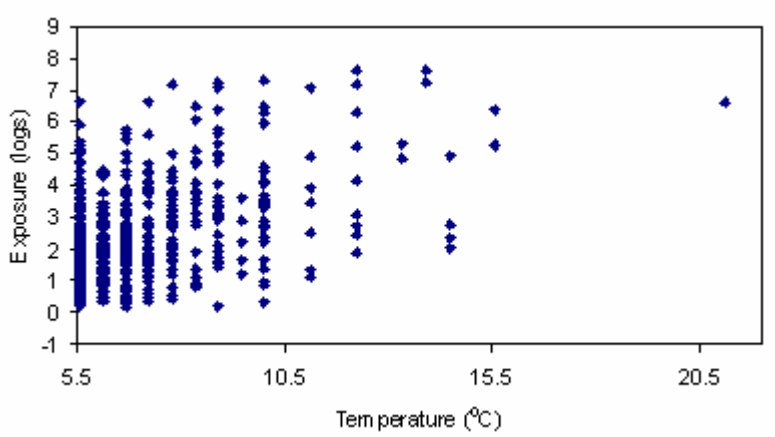
Figure 18-5. Scatter Plots of Exposure logs For Initial *Listeria monocytogenes* Concentration in Exposure Module of Milk Food category.




Lower Tail Region




Middle Region




Upper Tail Region

Figure 18-6. Scatter Plots of Exposure logs For Storage Temperature in Exposure Module of Milk Food Category.

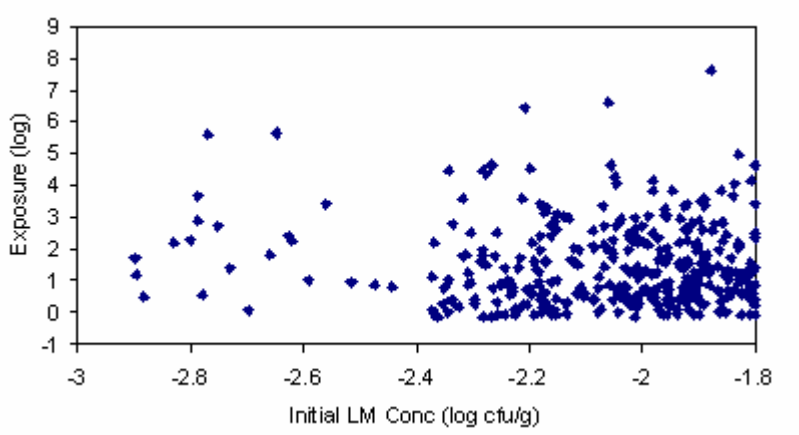
Regression Analysis, ANOVA and CART identified initial LM concentration and storage temperature as the top two important inputs for pâtés. The scatter plots for pâtés are shown in Figure 18-7 and 18-8 for inputs initial LM concentration and storage temperature, respectively.

The lower tail region scatter plots for initial LM concentration showed well spread response. There was no specific pattern in the response and it ranged from 0 logs to 8 logs. The middle region showed a linear increase in exposure value for initial LM concentration from -1.8 to 1.2 log cfu/g. The mean exposure values varied from 1.5 logs to 3 logs and the overall values ranged from 0 to 8 logs. The exposure values in upper tail region indicated saturation of exposure values around 4.5, 6 and 7.5 logs. The three levels corresponded to the maximum log growth at three different ranges of temperatures. An initial LM concentration of more than 3 logs always resulted in maximum log growth.

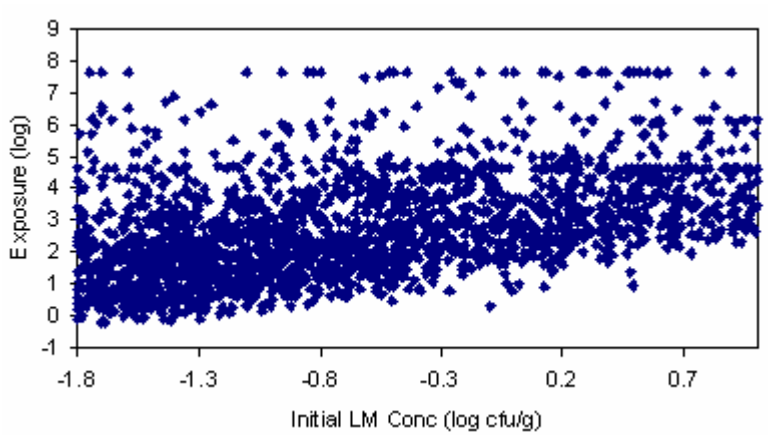
The scatter plots corresponding to storage temperature showed large variations in the exposure values. The exposure values ranged from -0.5 logs to 5 in lower and middle region of storage temperature. For storage temperature between 5 and 6.5 $^0$C the exposure values ranged from 0 to 6 logs. For storage temperature above 6.5 $^0$C the exposure values showed an increase in the mean response at point estimates of temperature. For storage temperature above 11 $^0$C the exposure values were always above 3.5 logs.

Regression Analysis, ANOVA and CART identified initial LM concentration and storage temperature as the top two important inputs for smoked seafood. The scatter plots for smoked seafood are shown in Figure 18-9 and 18-10 for inputs initial LM concentration and storage temperature, respectively.

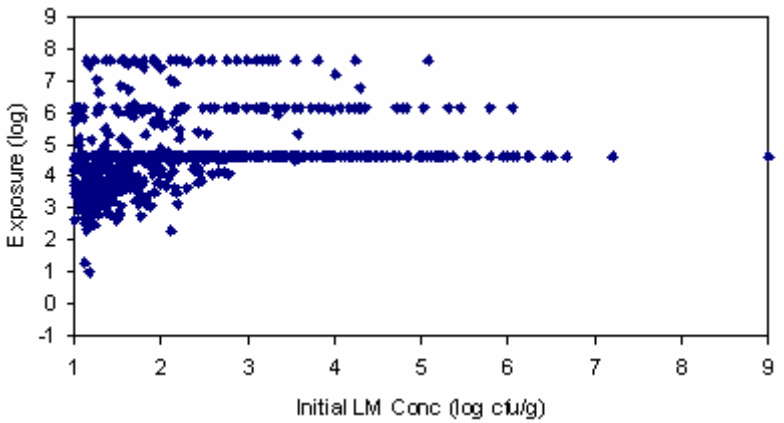
The lower tail scatter plots for initial LM concentration showed a linear increase from 0 to 1 log in mean response. Although most of the points lied between -0.5 logs to 2 logs, there were outliers as high as 8 logs. The middle region showed a linear increase in mean exposure values from 1 log to 3 logs and the overall exposure values ranged from -0.5 to 8 logs. The exposure values in upper tail region indicated saturation of exposure values around 4.5, 6 and 7.5 logs, for initial LM concentration value more than 3.5 logs. The exposure values increased linearly up to 3.5 logs. The storage temperatures versus exposure scatter plots did not indicate any specific pattern. For lower tail and middle region the exposure values lied between 0 to 5 logs for temperature up to 4.5 $^0$C and between 0 to 7 logs for temperature above 4.5 $^0$C. In the




Lower Tail Region

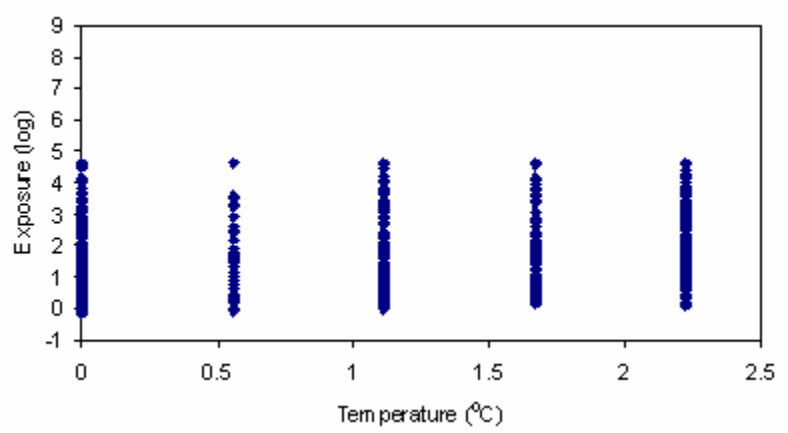



Middle Region

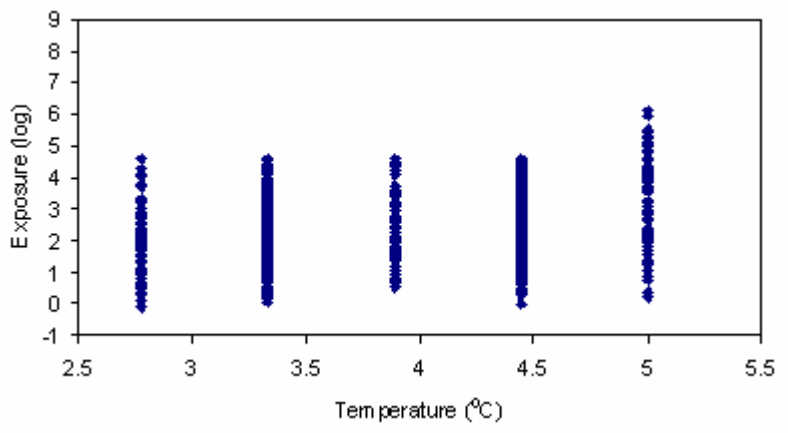



Upper Tail Region

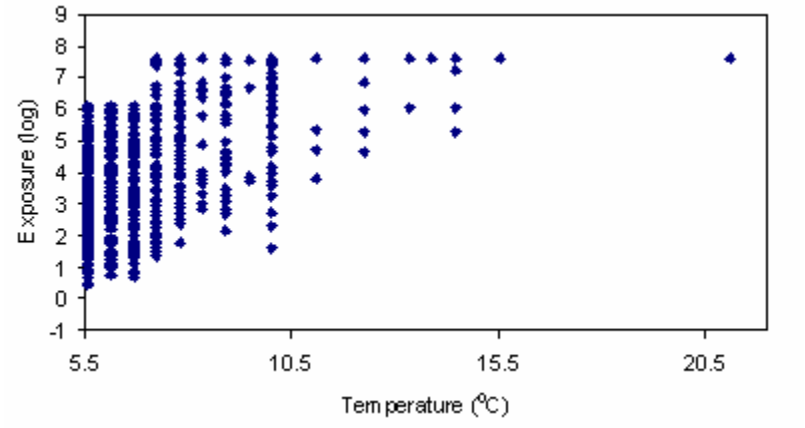
Figure 18-7. Scatter Plots of Exposure logs For Initial *Listeria monocytogenes* Concentration in Exposure Module of Pâté Food category.




Lower Tail Region

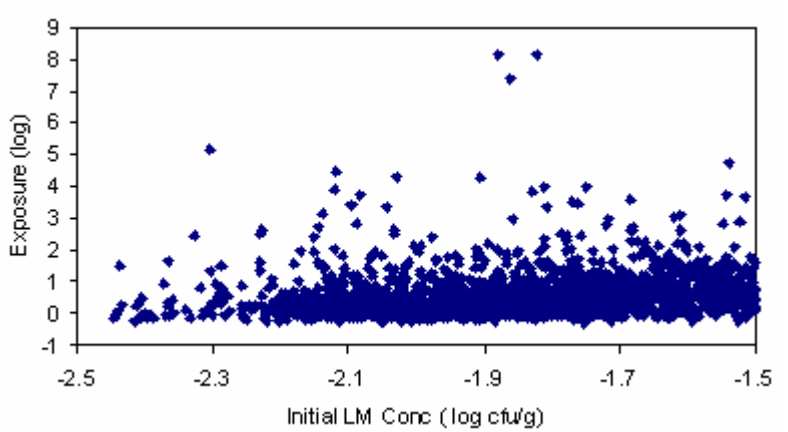



Middle Region

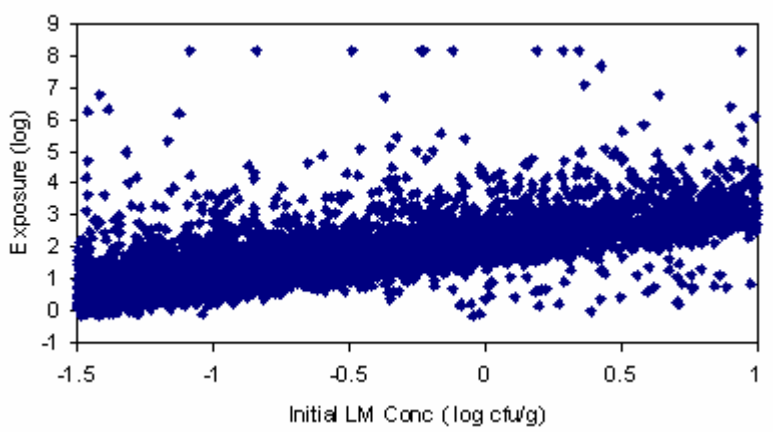



Upper Tail Region

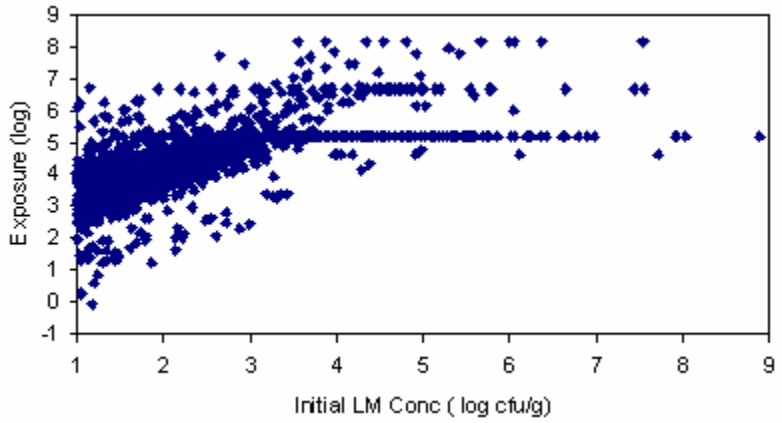
Figure 18-8. Scatter Plots of Exposure logs For Storage Temperature in Exposure Module of Pâté Food category.




Lower Tail Region

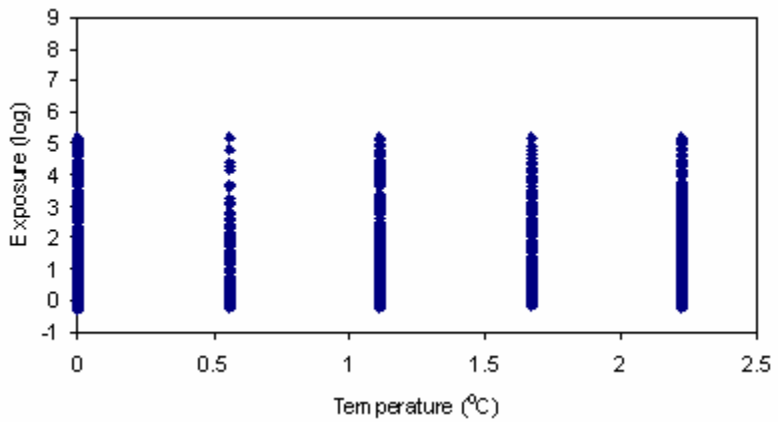



Middle Region

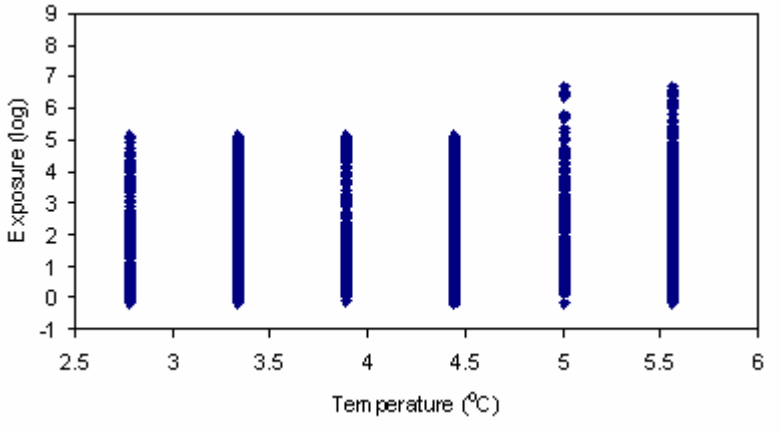



Upper Tail Region

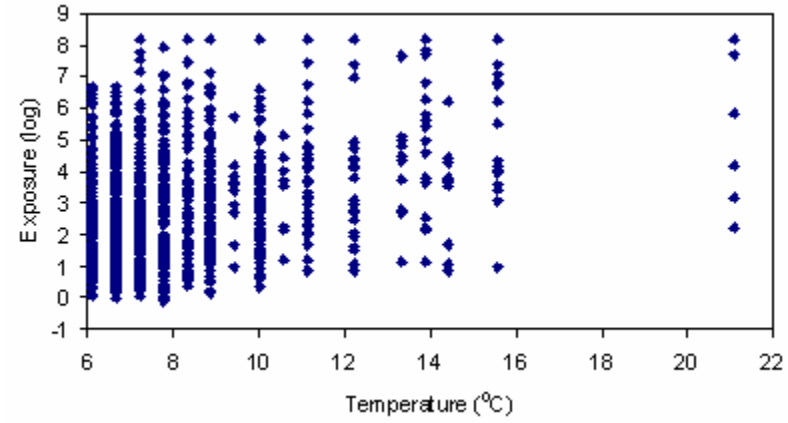
Figure 18-9. Scatter Plots of Exposure logs For Initial *Listeria monocytogenes* Concentration in Exposure Module of Smoked Seafood category.




Lower Tail Region




Middle Region




Upper Tail Region

Figure 18-10. Scatter Plots of Exposure logs For Storage Temperature in Exposure Module of Smoked Seafood category.

upper tail region the exposure values ranged from, 0 to 8.5 logs for temperature values greater than 6.7 $^0$C and 0 to 7 logs for temperature values less than 6.7 $^0$C.

The scatter plots in Figures 18-1 through 18-10 illustrate several key insights obtained from this approach to visualizing model behavior. In most cases, the trend of the average response of the output to the input was different for different ranges of the input. Thus, scatter plots provide insight into model responses that can be complex in various ways, such as because of threshold effects.

Scatter plots provide insight regarding the total amount of variability in the output for any given value of a particular input and regarding whether variation in an input contributes substantially to variation in the output. For example, if there is a large amount of scatter in the output for any given value of the input, and if the mean value of the output changes only modestly with respect to a change in input values, then it may be the case that the output is substantially sensitive to many other inputs aside from the one being evaluated.

The existence of saturation points or interactions with other inputs can be inferred from scatter plots. For example, the saturation of exposure with respect to the initial LM concentration was shown in Figure 18-7. Furthermore, the specific saturation level was temperature dependent. Scatter plots also can provide insight regarding the existence of changing upper or lower bounds for the output. For example, Figure 18-7 implies that the lower bound of exposure increases as the initial LM concentration increases, even though the upper bound of the simulated values remains approximately constant near the maximum possible saturation level.

Overall, scatter plots provide complex insight regarding the relationship between an output and selected inputs. These insights may be difficult or impossible to obtain via methods that do not include visualization of the database. Although it can be difficult to rank order key inputs using scatter plots, scatter plots can help confirm and clarify the effect of selected inputs with respect to the model output. Therefore, scatter plots can help provide insight regarding why and how a particular input is important.

## 18.2 Application of Scatter Plots to Dose Response Module

Scatter plots was applied to the dose response module for five selected food groups, including deli salad, fresh soft cheese, milk, pâtés, and smoked seafood. For each food group, the inputs of interest include uncertainty in dose adjustment factor, exposure period, virulence

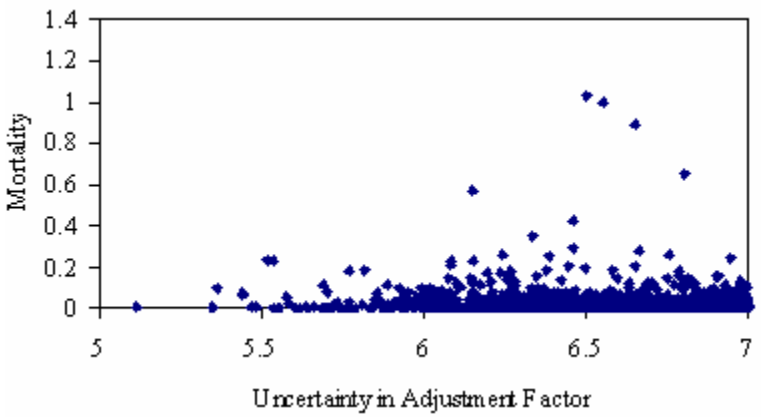
susceptibility uncertainty, and mouse lethality uncertainty and population exposure. The output of interest is mortality ascertained to the food group in the sub-population.

The dose response module incorporates only uncertainty. This section focuses on application of scatter plots based upon uncertainty in inputs. The total number of uncertain iteration applied was 4000. The adjustment factor was consistently ranked first with a few exceptions by previous sensitivity analysis methods and hence the direct dependency of mortality on adjustment factor was explored. The scatter plots for the deli salad, fresh soft cheese, milk, pâtés and smoked seafood categories are presented in Figures 18-11 to 18-15.

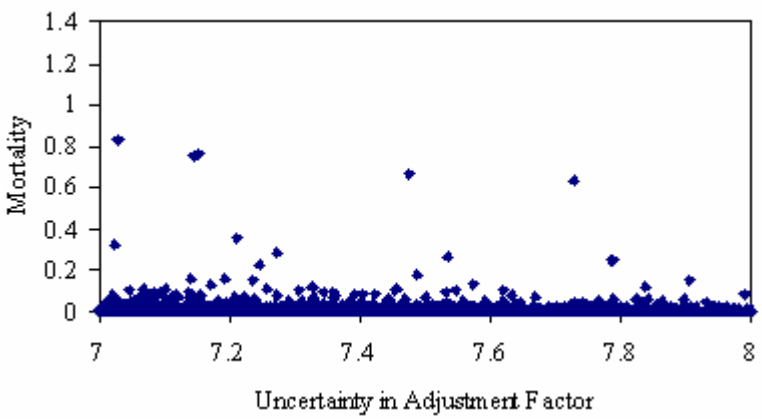
The scatter plots for deli salad are presented in Figure 18-11. The mortality values are mostly below 0.2 but in some cases is as high as 1.2. The variation in mortality for point estimates of adjustment factors decreased from lower tail region to upper tail region. The mortality values in lower tail region mostly ranged between 0 and 0.2 whereas in upper tail region were mostly between 0 and 0.1. The scatter plots for fresh soft cheese as shown in Figure 18-12 were similar to deli salad except that the mortality values were less by a factor of 10. The scatter plots for milk are shown in Figure 18-13. The variation in mortality corresponding to uncertainty in adjustment factor was big. The mortality values ranged from 0 to 14 in all the three regions. No specific pattern was identified. The scatter plots for pâtés and smoked seafood as shown in Figure 18-14 and 18-15 were similar to deli salad, but the ranges of mortality were from 0 to 4.5 and 0 to 7, respectively.

## 18.3 Summary of the Results of Application of Scatter Plots

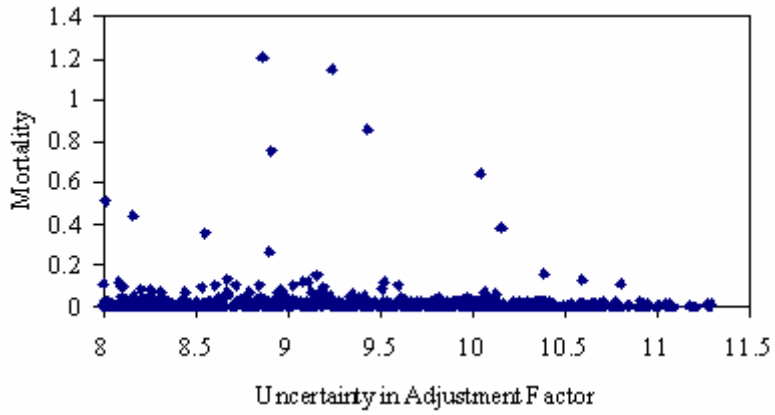
The application of scatter plots yielded insights about the response in exposure module. However the results of application of scatter plots on dose response module did not give much insight. The upper threshold or saturation level of exposure level was identified for fresh soft cheese, pâtés and smoked seafood. Also for all food categories the middle and upper tail regions of initial LM concentration showed linear increase in response. The scatter plots also showed that there was no direct dependency between storage temperature and exposure, as the ranges for point estimates of temperature was large.




Lower Tail Region

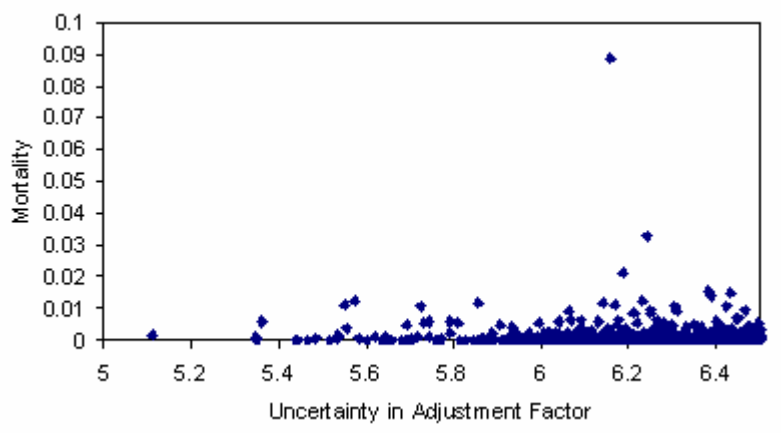



Middle Region

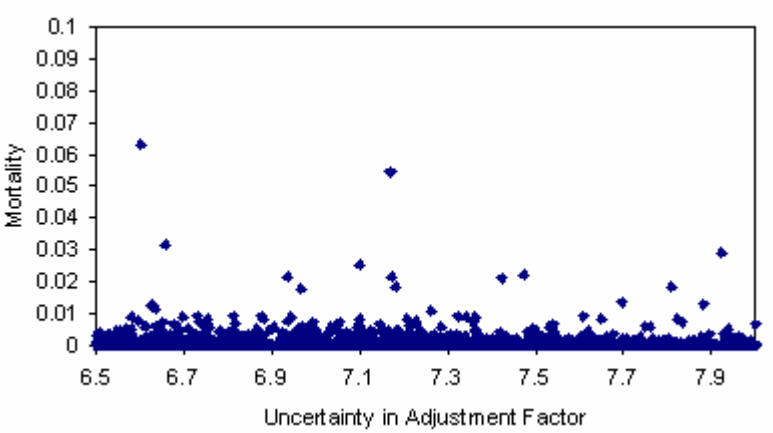



Upper Tail Region

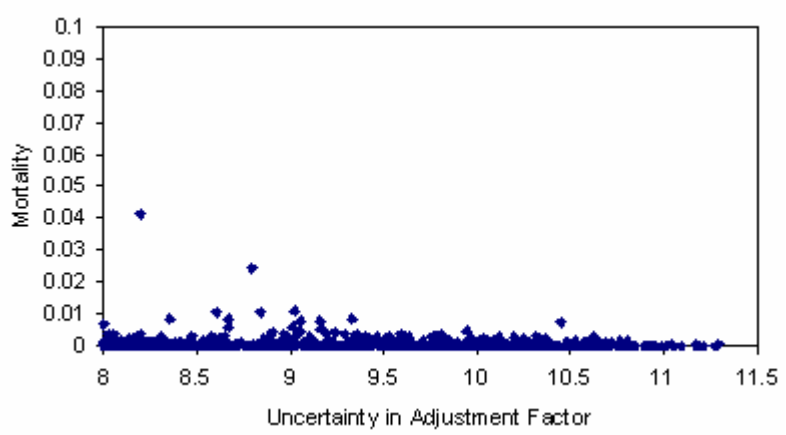
Figure 18-11. Scatter Plots of Exposure logs For Uncertainty in Adjustment Factor in Dose Response Module of Deli Salad category.




Lower Tail Region

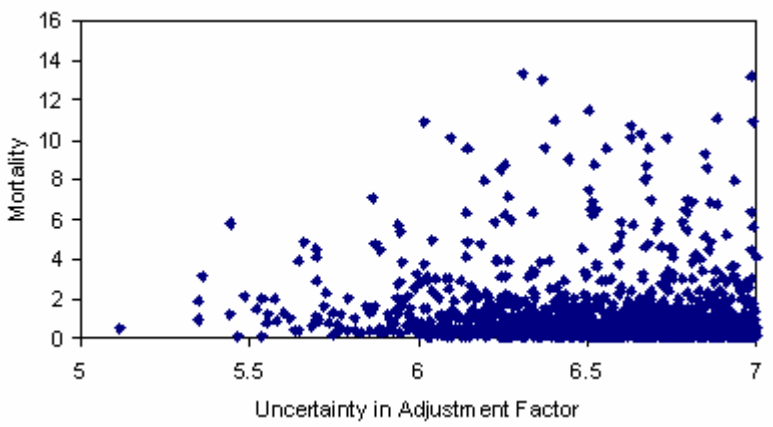



Middle Region

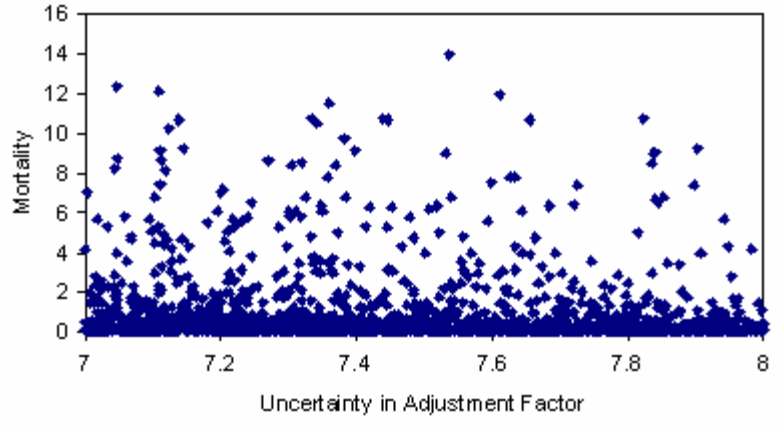



Upper Tail Region

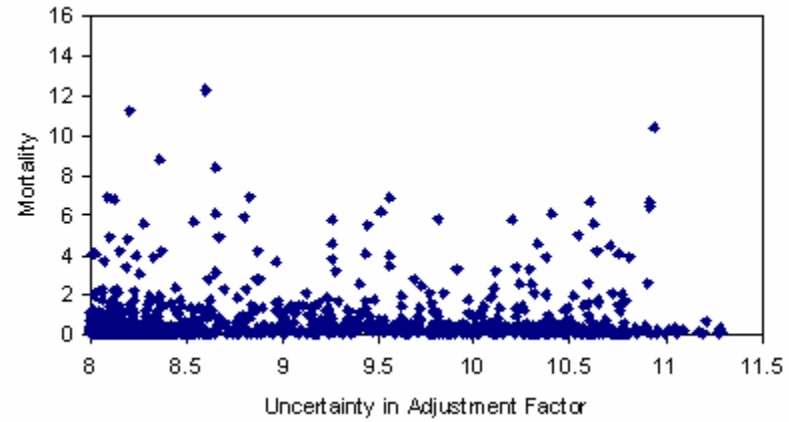
Figure 18-12. Scatter Plots of Exposure logs For Uncertainty in Adjustment Factor in Dose Response Module of Fresh Soft Cheese Food category.




Lower Tail Region

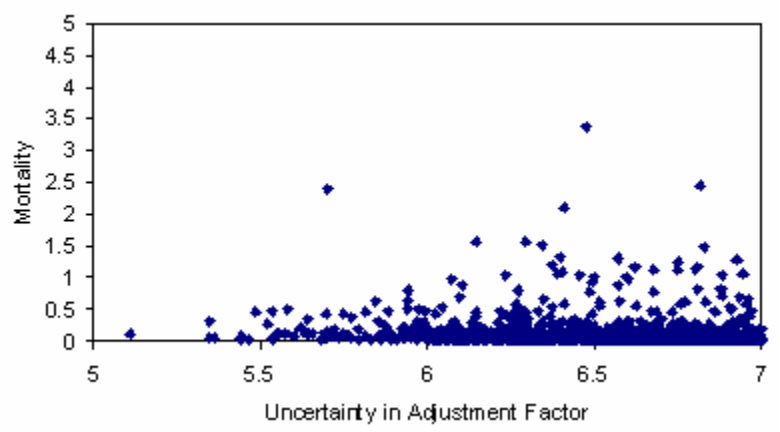



Middle Region

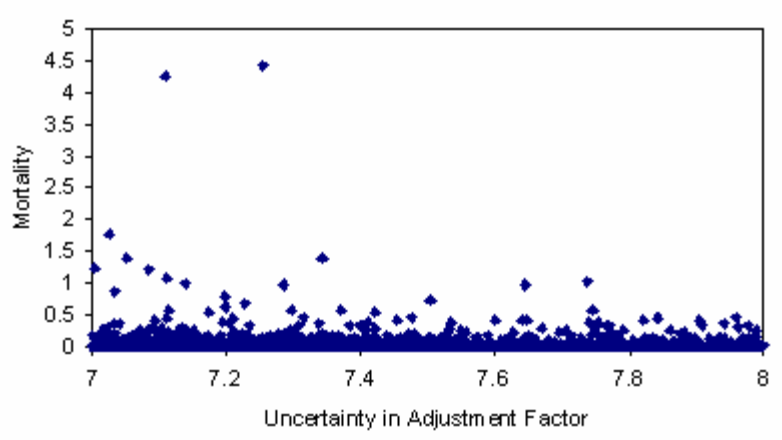



Upper Tail Region

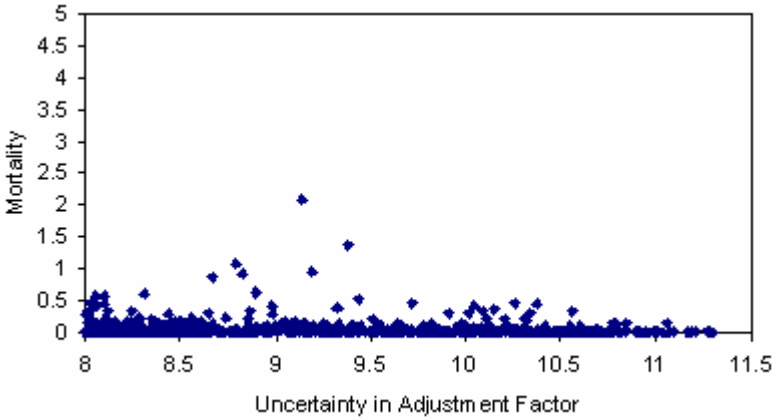
Figure 18-13. Scatter Plots of Exposure logs For Uncertainty in Adjustment Factor in Dose Response Module of Milk Food category..




Lower Tail Region

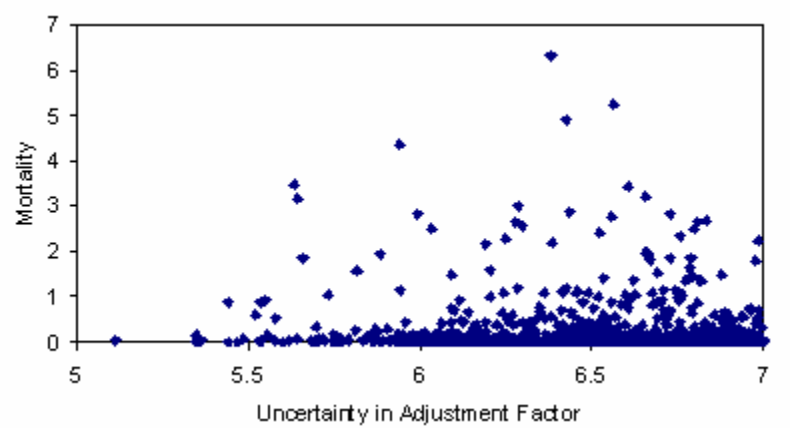



Middle Region

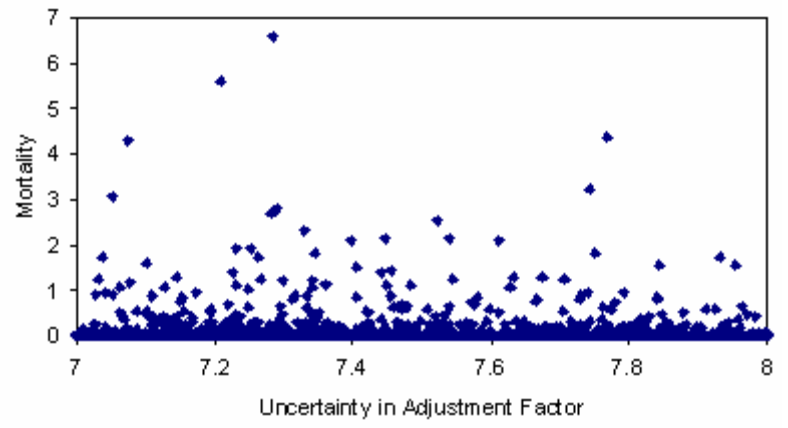



Upper Tail Region

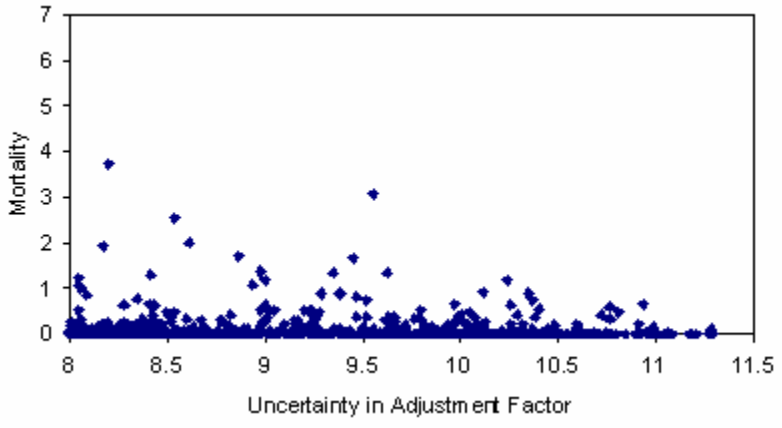
Figure 18-14. Scatter Plots of Exposure logs For Uncertainty in Adjustment Factor in Dose Response Module of Pâté Food category.




Upper Tail Region




Middle Region




Upper Tail Region

Figure 18-15. Scatter Plots of Exposure logs For Uncertainty in Adjustment Factor in Dose Response Module of Smoked Seafood category.

# 19 CONCLUSIONS BASED ON THE ANALYSES IN THE *LISTERIA MONOCYTOGENES* FOOD SAFETY RISK ASSESSMENT MODEL

In this chapter the results of the previous analyses on different modules and food categories of *Listeria monocytogenes* food safety risk assessment model are compared in order to evaluate selected sensitivity analysis methods.

In Chapters 13 to 18 different methods of the sensitivity analysis were implemented with the exposure and dose response modules of the *Listeria monocytogenes* food safety risk assessment model for five-selected food categories. The ranking of the inputs in each case are compared here for different methods. For example, the results are compared to see how different approaches of sensitivity analysis affect the ranking of the inputs in the model. This comparison is useful in order to have better insights regarding methods evaluation and the robustness of input rankings. Chapter 2 discussed the assumptions of different sensitivity analysis methods.

This section contains two parts. In the first part the results of the analyses performed in Chapters 13 to 18 on the exposure module for deli salad, fresh soft cheese, milk, pâtés and smoked seafood are discussed. In the second part the results of application of sensitivity analysis methods to the dose response module are compared.

## 19.1 Comparison of the Ranking of Inputs in the Exposure Module

The inputs to the exposure module were summarized in Section 12-7. There were five exposure modules corresponding to fresh soft cheese, deli salad, milk, pâtés and smoked seafood. Different methods of sensitivity analysis were applied on these parts as described in Chapters 13 to 18. These methods include NRSA, DSA, regression analysis, ANOVA, CART and scatter plot. In Tables 19-1 the rankings of the input based on methods of analysis that produce rankings are compared. A comparison is not included for scatter plots because scatter plots do not provide an indication of quantitative rankings. However, scatter plots provide insights into how the model responds to variations in an input. These insights are useful in explaining the results obtained with other methods.

With only a minor exception, all five of the sensitivity analysis methods identified the same top ranked input with regard to exposure associated with all five food categories. This agreement is notable in that it was obtained based upon methods with substantially different underpinnings. For example, NRSA evaluates only the individual effect of variation of a single input, not accounting for interactions among multiple inputs. DSA evaluates only the effect of a

Table 19-1. Comparison of Rankings Associated With Sensitivity Analysis Methods Applied to the Exposure Model Inputs for Deli Salad, Fresh Soft Cheese, Milk, Pâtés, and Smoked Seafood

| Input Variables | NRSA | DSA | REGRESSION[(a)] | ANOVA[(a)] | CART[(a)] |
|---|---|---|---|---|---|
| **Deli Salad** | | | | | |
| Serving Size (g) | 2 | 3 | 3,3 | 3,4 | 4[b], NA |
| Initial LM Concentration (log cfu/g) | 1 | 1 | 1,1 | 1,1 | 1, 1 |
| Storage temperature ($^0$C) | 3 | 2 | 4,4 | 4,3 | 3, NA |
| Storage time (days) | 5 | 4 | 5,5 | 5,5 | NA, NA |
| Growth at 5 $^0$C (log cfu/day) | 4 | 4 | 2,2 | 2,2 | 2, NA |
| **Fresh Soft Cheese** | | | | | |
| Serving Size (g) | 2 | 2 | 3,3 | 2,4 | NA, NA |
| Initial LM Concentration (log cfu/g) | 1 | 1 | 1,1 | 1,1 | 1, 1 |
| Storage temperature ($^0$C) | 3 | 3 | 4,5 | 4,3 | NA, 4 |
| Storage time (days) | 5 | 4 | 5,4 | 5,5 | NA, 3 |
| Growth at 5 $^0$C (log cfu/day) | 4 | 4 | 2,2 | 3,2 | 2, 2 |
| **Milk** | | | | | |
| Serving Size (g) | 2 | 5 | 5,5 | 4,5 | NA, NA |
| Initial LM Concentration (log cfu/g) | 1 | 1 | 1,1 | 1,1 | 1, 1 |
| Storage temperature ($^0$C) | 3 | 2 | 2,2 | 2,2 | 2, 2 |
| Storage time (days) | 4 | 3 | 4,3 | 5,3 | NA, 3 |
| Growth at 5 $^0$C (log cfu/day) | 5 | 3 | 3,4 | 3,4 | 3, NA |
| **Pâtés** | | | | | |
| Serving Size (g) | 2 | 5 | 5,5 | 5,5 | NA, NA |
| Initial LM Concentration (log cfu/g) | 1 | 1 | 2,1 | 1,1 | 1, 1 |
| Storage temperature ($^0$C) | 3 | 2 | 1,2 | 2,2 | 2, 2 |
| Storage time (days) | 5 | 3 | 3,4 | 3,4 | 3, NA |
| Growth at 5 $^0$C (log cfu/day) | 4 | 3 | 4,3 | 4,3 | NA, NA |

(Continued on next page.)

Table 19-1. Continued

| Input Variables | NRSA | DSA | REGRESSION[a] | ANOVA[a] | CART[b] |
|---|---|---|---|---|---|
| **Smoked Seafood** | | | | | |
| Serving Size (g) | 2 | 3 | 4,5 | 5,5 | NA, NA |
| Initial LM Concentration (log cfu/g) | 1 | 1 | 1,1 | 1,1 | 1, 1 |
| Storage temperature ($^0$C) | 3 | 2 | 2,2 | 2,2 | 2, 2 |
| Storage time (days) | 4 | 3 | 3,4 | 3,3 | NA, NA |
| Growth at 5 $^0$C (log cfu/day) | 5 | 3 | 5,3 | 4,4 | NA, NA |

(a) Two results presented: Variability only and co-mingled case.
(b) Two results are presented. The first is based upon visual inspection of the tree. The second is based upon the contribution to reduction in total deviance as a sensitivity index.

small perturbation in one input, not taking into account interactions. The three statistical methods of regression, ANOVA, and CART take into account simultaneous variation of all inputs according to their respective probability distributions. However, the specific variant of regression applied here was based upon a linearity assumption, whereas ANOVA and CART do not impose such an assumption regarding model form. Thus, it is remarkable that all five methods produce the same insight regarding the top ranked input. The agreement among the methods suggests that the model is sensitive to small perturbations of initial LM concentration, that the range of model response associated with the distribution of initial LM concentrations is larger than that associated with other inputs, even when all inputs are varying simultaneously, and that the model responds in approximately a linear manner to changes in initial LM concentration, without any substantial threshold effects.

The rankings for the second through fifth ranked inputs typically differ among the five sensitivity analysis methods, but often there are similarities within groups of methods. For example, the results of the three statistical methods often agree when differences in the sensitivity to inputs are statistically significant. In many cases in which the rank orders are reversed when comparing results with two of the statistical methods, it is because the two inputs in question are of comparable importance and, therefore, not statistically significantly different. The statistical methods typically agree regarding the least important input, or regarding the groups of inputs that have the lowest ranks. For example, for smoked seafood all three of the statistical methods agree regarding the first ranked input, the second ranked input, and the group

of three inputs for which there is little sensitivity. Variables that were not included in the regression trees were typically assigned low ranks by the other two statistical methods. Variables that were selected first in the regression trees were typically identified as the most important inputs based upon the other two statistical methods.

When comparing results for variability only and co-mingled variability and uncertainty with respect to the three statistical sensitivity analysis methods, there is typically agreement regarding the top ranked inputs and the least significant inputs. Apparent ambiguity regarding inputs of moderate to low importance is typically attributable to lack of statistical or practical significance of differences in sensitivity. For example, for deli salads, the first, second, and fifth ranks are assigned to the same inputs for all three statistical methods for both probabilistic analyses. The apparent differences in the assignment of the third and fourth ranks when comparing variability only results among the methods, or when comparing the method results for two different probabilistic simulations, are attributable to a lack of substantial difference in the sensitivity to the serving size and storage temperature.

Although the results from NRSA and DSA agree with those of the other methods for the top ranked input, they differ from each other in minor ways and from the statistical methods in more significant ways with regard to the ranking of other inputs. For example, the lowest ranked input obtained using NRSA is not always the lowest ranked input obtained from the statistical methods. This difference in results suggests some importance associated with simultaneous variation in multiple inputs that is accounted for in the statistical methods but not NRSA. The input ranked lowest by DSA is typically also ranked low by NRSA but on occasion is ranked substantially higher by the statistical methods. For example, DSA leads to a last place tie in the ranking of growth at 5 $^{0}$C, whereas this input is ranked second by the statistical methods. The difference in ranking in this case is attributed to DSA's failure to consider simultaneous interactions among inputs. Although it is also the case that DSA considers only a small perturbation of the input, the low ranking for growth at 5 $^{0}$C was also obtained using NRSA, which considers a wider range of variation. DSA and NRSA occasionally disagree regarding identification of the second most important input.

Overall, the comparison of results with five sensitivity analysis methods, including in some cases two different probabilistic simulations, suggest that it is often possible to identify the most and least sensitive inputs with a fair degree of reliability regardless of the choice of method

as long as the model is approximately linear and does not include a substantial threshold effect. However, a robust identification of inputs of secondary importance will depend more critically on the selection of an appropriate method. These results are specific to the model, but suggest that the statistical methods will provide results consistent with the linear-based mathematical methods while also including a capability, in the cases of ANOVA and CART, to respond more appropriately to nonlinearity or threshold effects if present.

## 19.2 Comparison of the Inputs Ranking in the Dose Response Module of the Five Food Categories

The inputs to the dose module were summarized in Section 12-7. There were five dose response modules corresponding to fresh soft cheese, deli salad, milk, pâtés and smoked seafood. Different methods of the sensitivity analysis were applied to these food categories as described in Chapters 13 to 18. In Table 19-2 the rankings of the inputs based on five of the methods of sensitivity analysis are compared.

The comparison of results for the dose-response module is more complex than for the exposure module. However, the comparison can be made separately for the two mathematical methods and for the three statistical methods. For all five food groups, the two mathematical methods provided the same ranking of the three least important inputs. The two mathematical methods provided reversed rankings of the top two inputs. The top two inputs as identified using the mathematical methods occasionally, but for the most part did not, correspond to the top inputs identified using the statistical methods. For example, the fraction of population exposed and the mouse lethality were typically the two most important inputs identified using the mathematical methods. In contrast, the dose adjustment factor, the exposure period, and to a lesser extent, the fraction of population exposed, were among the top two most important inputs for a given food group based upon the statistical methods. Thus, mouse lethality was not identified as one of the most important inputs based upon the statistical analysis results.

The similarities in results between NRSA and DSA suggest that the ranges of values used in NRSA were typically proportional to the ranges used for perturbations in DSA. However, since neither methods account for simultaneous variation among the inputs, it is expected that there could be differences when compared with results from the statistical methods.

Table 19-2. Comparison of Rankings Associated With Sensitivity Analysis Methods Applied to the Dose Response Model Inputs for Deli Salad, Fresh Soft Cheese, Milk, Pâtés, and Smoked Seafood

| Input Variables | NRSA | DSA | REGRESSION | ANOVA | CART |
|---|---|---|---|---|---|
| **Deli Salad** | | | | | |
| Dose Adjustment Factor | 5 | 5 | 2 | 1 | 2 |
| Exposure Period | 4 | 4 | 3 | NA | NA |
| Virulence Susceptibility | 3 | 3 | NA | NA | NA |
| Mouse Lethality | 1 | 2 | NA | NA | NA |
| Fraction of Population Exposed | 2 | 1 | 1 | NA | 1 |
| **Fresh Soft Cheese** | | | | | |
| Dose Adjustment Factor | 5 | 5 | 1 | 1 | 2 |
| Exposure Period | 4 | 4 | NA | NA | NA |
| Virulence Susceptibility | 3 | 3 | NA | NA | NA |
| Mouse Lethality | 2 | 2 | NA | NA | NA |
| Fraction of Population Exposed | 1 | 1 | NA | 2 | 1 |
| **Milk** | | | | | |
| Dose Adjustment Factor | 5 | 5 | 1 | 1 | 1 |
| Exposure Period | 4 | 4 | 2 | 2 | NA |
| Virulence Susceptibility | 3 | 3 | NA | NA | NA |
| Mouse Lethality | 2 | 2 | NA | NA | NA |
| Fraction of Population Exposed | 1 | 1 | NA | NA | 2 |
| **Pâtés** | | | | | |
| Dose Adjustment Factor | 5 | 5 | 1 | 1 | 1 |
| Exposure Period | 4 | 4 | 2 | NA | 2 |
| Virulence Susceptibility | 3 | 3 | NA | NA | NA |
| Mouse Lethality | 1 | 2 | NA | NA | 3 |
| Fraction of Population Exposed | 2 | 1 | NA | NA | NA |

(Continued on next page.)

Table 19-2. Continued

| Input Variables | NRSA | DSA | REGRESSION | ANOVA | CART |
|---|---|---|---|---|---|
| **Smoked Seafood** | | | | | |
| Dose Adjustment Factor | 5 | 5 | 1 | 2 | 2 |
| Exposure Period | 4 | 4 | 2 | NA | NA |
| Virulence Susceptibility | 3 | 3 | NA | NA | NA |
| Mouse Lethality | 1 | 2 | NA | NA | NA |
| Fraction of Population Exposed | 2 | 1 | NA | 1 | 1 |

The three statistical methods typically agreed, with only a minor exception, that virulence susceptibility and mouse lethality were not significant. However, the results for other inputs vary among the five food categories. For example, although exposure period was often not identified as significant based upon ANOVA and CART, for four of the five food groups it was identified as significant based upon regression analysis. Since the form of regression used here differs from ANOVA and CART, most noticeably with respect to a linearity assumption and failure to directly account for threshold effects, it is possible that an interaction effect, nonlinearity in response, or a threshold effect could be responsible for the difference in rankings.

The dose adjustment factor was typically identified as either the first or second ranked input among the five food categories based upon the statistical methods, whereas it was identified as least important based upon the two mathematical methods.

The differences in results when comparing the statistical versus mathematical methods suggest the importance of accounting for simultaneous variation in the inputs. Approximate similarities in results among the statistical methods suggest that the model is not strongly nonlinear and that it does not have a strong threshold effect. However, as noted in Chapter 15, the $R^2$ values for regression applied to the dose response module were low, which suggest that a linear model is not a good fit to the probabilistic simulation results. In combination, these results suggest that if there are substantial differences in sensitivity among the inputs, it can be possible to obtain an accurate identification of the top ranked input using regression analysis even though the goodness-of-fit is poor. Of course, it is not recommended to rely upon regression results if the goodness-of-fit is poor. In such cases, it is prudent to make a comparison with another method, or to use another method instead. Although the identification of the top ranked input

appears to be robust to the choice of method for this particular model, the identification of the second ranked input is more sensitive to the selection of an appropriate sensitivity analysis method.

# 20 CONCLUSIONS AND RECOMMENDATIONS

The objective of this chapter is to present the conclusions and give recommendations based upon the application of sensitivity analysis methods to the *E. coli* and *Listeria monocytogenes* models. This chapter addresses the questions raised in Chapter 1 based on analysis results in Chapters 4 to 19.

This chapter contains two parts. Section 20.1 answers the key questions raised in Section 1.1. It discusses the advantages and disadvantages for application of each sensitivity analysis method based on this study. Section 20.2 presents the recommendations for application of sensitivity analysis methods to food safety risk assessment models.

## 20.1 Key Criteria, Advantages and Disadvantages for Application of Different Sensitivity Analysis Methods

Nine methods of sensitivity analysis were applied to the *E. coli* and *Listeria monocytogenes* models. These methods include NRSA, DSA, standardized linear regression analysis, rank regression, correlation coefficients, ANOVA, CART analysis, scatter plots and conditional sensitivity analysis. In some cases, variations of specific methods were evaluated, such as comparisons of Pearson (sample) and Spearman (rank) correlation, and comparisons of linear regression analysis using alternative nonlinear basis functions. The NCSU/USDA Workshop on Sensitivity Analysis, 2001, listed the characteristics of the food safety risk assessment models that an ideal sensitivity analysis method would seek to address. These characteristics are:

- nonlinearities in response;
- thresholds (e.g., below which there is no growth of a microbial pathogen);
- interaction between inputs (e.g., low storage temperature is often associated with high storage time );
- qualitative and quantitative inputs;
- ability to identify and characterize high exposure cases; and
- ability to deal appropriately with variability and uncertainty, such as via two dimensional probabilistic analysis.

The analyses of this work were targeted to answer several key questions based on these characteristics that address the overall project objective. These key questions, as raised in Section 1.1, are:

- Can simple sensitivity analysis methods such as nominal range sensitivity analysis provide robust insights in spite of their apparent limitations?
- Which methods can take care of qualitative and quantitative variables simultaneously?
- Which methods can identify and appropriately respond to thresholds?
- Which methods can specifically address high exposure/risk case scenarios?
- Which methods can give insights on interactions between explanatory variables?
- Which methods can identify or appropriately deal with non-linearity in response?
- How unambiguous is the relative importance of the model inputs based on the selected sensitivity index?
- How should sensitivity analysis be conducted in a two-dimensional probabilistic framework?
- Which sensitivity analysis methods can be easily automated to address the additional complexity introduced by two-dimensional Monte Carlo simulation of variability and uncertainty?

The answers to these questions, advantages and disadvantages of each method as identified by the study are discussed for each of the methods in Sections 20.1.1 to 20.1.8. Section 20.1.9 summarizes the key characteristics addressed by each of the sensitivity analysis methods.

It is important to understand that the focus of this study was to evaluate sensitivity analysis methods based on the ability to address the above questions. In practical applications, sensitivity analysis methods would be used to gain insights with respect to specific decision objectives. There was not a specific decision context for the development and application of the *E. coli* model. Decision-based analysis can give different results for different objectives. For example, in the *E. coli* model, cooking part was the most important module based on maximum per serving risk whereas was most important for maximum number of contaminated serving. In cases where the objective of application of sensitivity analysis is to aid in making research and management decisions, the selection and application of methods must address the decision focus.

### 20.1.1 NRSA

NRSA identified the most important input consistently. NRSA is a relatively simple and easy to apply method. The number of model evaluations required for NRSA is proportional to the number of inputs for which sensitivity analysis is needed. NRSA provides a numerical measure of sensitivity. NRSA is based upon an implicit linear assumption regarding the model

form. Because NRSA does not consider simultaneous variation of all inputs, NRSA cannot deal with interactions between inputs that may arise because of nonlinearity, nor may it lead to insight regarding the likely range of variation in the model output. NRSA does not provide clear insight regarding whether thresholds exist or for what part of the domain of model inputs that they apply. There is no measure of statistical or other significance for the sensitivity coefficients by which to make a clear judgment of whether one input is substantially or significantly different in importance from another.

NRSA can be applied to both qualitative and quantitative inputs. However, thresholds cannot be identified by NRSA as it calculates ranking based upon point estimates and not distributions of inputs. A threshold can lie anywhere within the range of an acceptable values of an input or may not be realized if the threshold depends on interaction with other inputs.

NRSA varies only one input at a time keeping others at point estimates. This approach implicitly assumes that the model is linear. If the model is nonlinear, there could be specific combinations of inputs that result in higher exposure scenarios than would be the case for a linear model. Therefore, NRSA might fail to capture in an appropriate manner how the model would respond if variation in the inputs was considered simultaneously.

In cases where interactions between inputs are present in the model, the response of the model to a point estimate of one input depends on the point estimates of another input. In calculation of sensitivity indices using conventional NRSA, only one variable is varied at a time. This restriction eliminates the possibility of identifying any interaction between inputs associated with nonlinearity of the model. However, conditional NRSA, in which NRSA is performed multiple times based upon different point estimates, can provide insight regarding interactions. The application of conditional NRSA is more time consuming than for conventional NRSA. Furthermore, it can be difficult to develop a clear ranking among all inputs if the results differ conditioned on one point estimate of the inputs versus results conditioned on a different point estimate.

Thus, NRSA has important conceptual limitations that suggest that its performance is not likely to be robust when applied to models with nonlinearities and thresholds. In practice, based upon the case study results, NRSA performed adequately in correctly identifying the most important input, but performed inconsistently in identifying other significant inputs that were of less importance.

### 20.1.2 DSA

DSA addresses the effect of small variations around a point estimate of an input. Therefore, DSA is a method for evaluating how sensitive the model response is to small perturbations in an individual input when all other inputs are kept constant. In this regard, DSA differs from NRSA in that only small perturbations, and not the likely ranges of variation, are considered for each input. Thus, although DSA can provide useful insight regarding how the output responds to a unit change in the input, DSA does not provide insight regarding how the output responds to a full range of variation of the input. In other respects, DSA is similar to NRSA and suffers from the same limitations. However, unlike NRSA, it is more difficult to work with qualitative inputs with DSA because for such inputs there is typically not a meaningful way in which to perturb them by one unit or by an arbitrarily small amount.

### 20.1.3 Regression Analysis

Regression analysis assumes a functional form of the mathematical model and measures its goodness of fit. It is possible to use linear combinations of nonlinear basis functions within the framework of linear regression analysis, or to use nonlinear regression instead. However, the effort involved to find a suitable functional form is typically based upon an iterative process that can be time consuming. Moreover, it is possible to use rank rather than the sample regression. However, linear sample regression is commonly used technique. Therefore, for simplicity and clarity, the objective of this work was to narrowly focus on how well linear standardized sample regression performs when applied consistently over different modules and parts of the two selected food safety risk models. In many cases the rankings obtained with linear regression analysis were comparable to those of other methods even when the goodness of fit was poor.

Rank regression is an alternative regression-based approach applicable for cases subject to non-linearity in the model response. Hence, for cases with low $R^2$ values based on the linear regression analysis, rank regression may provide better regression fit to the rank-based dataset capturing higher amount of output variation. Rank regression is appropriate when there is a monotonic response in the model.

The *Listeria monocytogenes* model did not have qualitative inputs whereas *E. coli* model had both qualitative and quantitative inputs. In the *E. coli* model, a few cases were presented in which some inputs were qualitative. These inputs were addressed in regression analysis using indicator variables instead of the original qualitative inputs. In such cases, coefficients were

estimated for indicator variables representing the mean model responses at each qualitative input level. There were no unique regression coefficient estimates for the qualitative inputs. F values are estimable for both qualitative and quantitative inputs. Hence, instead of regression coefficients, F values were used as a sensitivity index when there were qualitative inputs to the model. In general, the ranking of the importance of inputs using F values was reasonably consistent with rankings or insights obtained from other methods, including graphical approaches.

Because of the linear functional relation assumption between the output and inputs that is the basis for linear standardized sample regression; inferences cannot easily be made regarding the presence and effect of thresholds on the regression results. These types of model responses may give rise to low values of the coefficient of determination when linear regression is used. However, other regression analysis techniques such as the change-point regression method, where the parameters to be estimated are the points at which the slopes change, may be able to identify thresholds (Ogden and Parzen, 1996).

The specific ranges of values for key variables that lead to high exposure cases cannot be specifically identified based on regression analysis. Regression analysis does not easily permit distinction between high and low exposure cases, since the complete range of inputs are considered as part of a single analysis.

For quantitative inputs, 95 percentile confidence intervals were provided for estimated regression coefficients to give insight regarding how unambiguous the ranks are based on the relative magnitude of the coefficients as the sensitivity index. Confidence intervals consider both the magnitude of the regression coefficient and the associated error to the estimated coefficients. Hence, confidence intervals represent the amount of uncertainty in the estimated regression coefficients. The output presents comparable sensitivity to inputs with overlapping confidence intervals.

For the applications of regression to the *Listeria monocytogenes* and *E. coli* model, three probabilistic simulations were implemented when both variability and uncertainty were considered: (1) variability only; (2) variability for different uncertainty realizations; and (3) one dimensional co-mingled variability and uncertainty. The comparisons of rankings from these methods provide insights about the robustness of rankings of an input. The most important input was the same for all three simulations when applying regression analysis. There was agreement

on a group of secondary importance inputs. In some cases, the inputs grouped as moderate importance or least importance variables did not show a complete agreement based on all three simulations. However, since it is typically more important from a decision making and risk management strategy to robustly identify the most important inputs, the similarity of results among the three probabilistic simulation approaches suggests that the insights are robust regardless of which approach is employed. However, these results merely suggest but do not prove that this is the case in general, since these results are based upon analysis with a particular model and not with all possible models.

### 20.1.4 Correlation Coefficients

Correlation coefficients measure the strength of a linear association between an input and the output. Two types of correlation coefficients were evaluated for sensitivity analysis, including Pearson (sample) and Spearman (rank) correlation coefficients. Sample and rank correlations are commonly used by practitioners because these methods are often included in commercial software packages, such as Crystal Ball, that are used in many risk assessment studies.

The sample correlation coefficients method assumes a linear association between the output and individual inputs. Hence, application of this method for sensitivity analysis does not provide reliable results when there is non-linearity inherent in the model. The rank correlation coefficients method is recommended in cases with non-linearity in the model as this method does not assume any specific functional association between the output and the inputs. Nonetheless, rank correlation coefficients method requires a monotonic response in the model.

Both sample and rank correlation coefficients do not provide any statistical measure for identifying possible threshold in the model response. Interaction effects between inputs, as one of the characteristics of the food safety models, cannot be captured using coefficient methods, nor can methods identify the ranges of inputs associated with low or high exposure scenarios.

Inputs were ranked based on the relative magnitude of the estimated correlation coefficients. Sample and rank correlation coefficients were presented for the two-dimensional simulation of the variability under several uncertainty realizations for a selected part of the *E. coli* model. Application of these methods to two-dimensional frameworks can provide insights regarding how uncertainty in the inputs can affect the estimates of the relative ranks of the inputs with regard to variability or vice versa. The macro feature in SAS was implemented for

automation of the application of these methods to the two-dimensional simulation of variability and uncertainty and summarization of the results.

### 20.1.5 ANOVA

ANOVA is a generalized probabilistic sensitivity analysis method used to determine whether there is a statistically significant association between an output and one or more inputs. In those cases that the $R^2$ value for a linear regression analysis indicates that the linear assumption for the model is not valid, ANOVA can provide a more reliable ranking. Furthermore, categorical inputs, groups of inputs, and interactions can be addressed. ANOVA requires quantitative inputs to be categorized into levels. The *E. coli* model is an example of a case where both qualitative and quantitative inputs were part of the model.

Contrasts in ANOVA were presented as a tool to identify thresholds in the model response to an input. Identification of thresholds using ANOVA depends on the method used for defining levels for each input. Contrasts in ANOVA can also be used to identify the non-linearity in the model response to each of the inputs.

In this report three methods were presented for *E. coli* model in order to define levels for each input: (1) equal intervals; (2) equal percentiles; and (3) visual inspection of the CDF for each input. The identification of thresholds is facilitated using equal intervals. For example, Section 5.4.1.3 presented a case that contrasts were used to identify a threshold in the growth of the *E. coli* organisms to the storage time. The levels for the *Listeria monocytogenes* model were created using percentiles as described in Section 12.9. The levels considered corresponded to 0 to 0.2, 0.2 to 0.8 and 0.8 to 1.0 percentiles. These percentiles were assumed to represent lower tail, middle region and upper tail of the cumulative distribution plot for the input. In general, it is preferable to define levels of an input so that each level is comparatively homogeneous and does not include a threshold. However, it is also preferable to define levels so that each level has an adequate sample size and so that the stratification of the inputs is "balanced," as discussed in Chapter 2, with respect to the estimation of contrasts.

ANOVA is a powerful sensitivity analysis method for purposes of identifying interactions among inputs. For factorial experiments with contributions of more than one input in the model in addition to the simple effect of each input, the interaction between inputs should be considered in ANOVA. The interaction terms can be compared to each other based on the magnitude of their estimated F values. Using equal intervals for input levels makes it possible to

ascertain any statistically significant functional relation between the output and an input by considering appropriate contrast.

The uncertainty in point estimates of F values, which are used as a sensitivity index in ANOVA, should be taken into account when making comparisons of the F values of two or more factors. The evaluation of uncertainty in F values provides insight regarding the ambiguity of ranks based on F values. The method of bootstrap simulation was used to generate sampling distributions of uncertainty for F values. The sampling distribution of F values for the factors can be compared to infer whether there is ambiguity in the relative ranking. In situations where the F values are similar, factors can be categorized into groups of similar importance. It is also possible to discriminate between groups of factors such that there are clear differences between groups. The results will differ for sample size other than those used here and for other models and case studies.

As explained in Section 20.1.3, three probabilistic simulations were implemented for variability only, co-mingled variability and uncertainty, and variability under uncertainty. The comparisons of rankings from these simulations provide insights regarding the robustness of rankings of an input. The most important input was the same for all three simulations when applying ANOVA. There was also agreement on a group of inputs of secondary importance. In some cases, the inputs grouped as of moderate importance or least importance variables did not show a complete agreement based on all three simulations. Thus, similar to the insights obtained regarding regression analysis as described in the previous section, the insights regarding the key inputs were typically robust regardless of which approach probabilistic simulation approach was employed.

ANOVA was not used to demonstrate two-dimensional analysis of variability under uncertainty for *Listeria monocytogenes* model, although the macro feature in SAS can be used to automate application of ANOVA and to summarize results. ANOVA requires that each input must be divided into levels. The levels in the case studies presented were created based upon percentiles of the input distribution. However, in a two dimensional analysis, the levels should remain the same across all uncertainty simulations to enable comparison of rankings for variability under different uncertainty cases. In such cases, the number of data points within each treatment for uncertainty iterations were found to be highly unbalanced. Thus, the results of the two dimensional analyses if applied in this manner would not be reliable. However, it is possible

that for other models and case studies the data may be less unbalanced, making application of two-dimensional analysis possible. Thus, depending on the case study, the results of the two dimensional analyses are found to be robust only if the data points within each treatment are not extremely unbalanced. The case studies provided for *E. coli* show that the results can be robust in spite of slightly unbalanced data. However, two dimensional analyses is not advisable for highly unbalanced treatments.

### 20.1.6 CART

CART analysis does not assume any specific model and classifies data into a number of groups. The data within a class are more homogeneous in response than the total dataset. CART was identified as a powerful method for addressing both qualitative and quantitative inputs without any pre-processing of the dataset. Moreover, CART does not assume a specific functional relation between the model inputs and the model response. Hence, for models that have nonlinearity or thresholds, the application of CART does not force any underestimation or overestimation regarding the sensitivity of the output to each input.

The ranking of the inputs in CART is typically based on visualization of the regression tree and judgment. In some cases, the application of other sensitivity analysis methods as a complement to CART was needed to gain insight regarding the rank of input that were of secondary importance. Because CART does not automatically produce a sensitivity index similar to those of methods such as ANOVA or regression analysis, it was difficult to apply CART in repetitive frameworks. For example, when variability was simulated separately for multiple realizations of uncertainty in a two-dimensional probabilistic framework, it was difficult to summarize and compare the results of the CART analysis for each of the uncertainty realizations. Thus, the lack of a quantitative sensitivity index is a practical limitation that makes it difficult to automate CART for use with two-dimensional probabilistic analysis. However, it would be possible to apply CART in a repetitive framework in a simplified manner, such as by recording which input was selected as the basis for the first split and which input(s) was (were) selected as the basis for the second level of splits in the tree.

Although there is no direct sensitivity parameter in CART, a possible sensitivity index was explored in selected case studies with both *E. coli* and *Listeria monocytogenes* models. Specifically, the sum of the reduction in deviance by each input included in the tree was used as a sensitivity index. Hence, inputs were ranked based on the percentage of their contribution to

the reduction of the total deviance. However, this parameter is not a direct output from the S-PLUS GUI interface. Manual calculation of the sensitivity index is tedious. The development of software to calculate a sensitivity index is recommended.

The results obtained for ranks of inputs from the inspection of regression tree mostly agreed with the ranks obtained by using the contribution to reduction in deviance as an alternative sensitivity index. However, in some cases it was difficult to comment on the ranks of the inputs just based on the inspection of regression tree. Alternative sensitivity indices were useful in assigning relative ranks to each of the inputs that participated in the tree. Thus, in cases where it is difficult to decide ranks based upon just inspection of the tree, the use of alternative sensitivity index is highly desirable.

CART analysis is expected to respond appropriately to the existence of thresholds in the model response. However, it may not be obvious simply by inspection of CART results as to whether the points used to partition a variable represent an actual threshold or are merely a convenient point of division for a variable that is varying in a continuous manner. In order to verify whether the condition specified in a node represents a threshold in the model response, results from other sensitivity analysis methods were used, including graphical methods.

A key advantage of CART compared to all other mathematical and statistical methods that were evaluated is that CART provides explicit insight into the ranges of values of key inputs that lead to the worst, or best, outcomes. For example, the classification rules for the highest mean exposure identify the combination of ranges of key inputs that lead to the highest model output. The identification of highest exposure classes is important to identifying the possible approaches to mitigate exposure and/or risk.

### 20.1.7 Scatter Plots

Scatter plots were used for visual assessment of the influence of individual inputs on an output. Scatter plots are often recommended as a first step in sensitivity analysis of a statistical sample of data. A key advantage of scatter plots is that they allow for the identification of potentially complex dependencies that cannot easily be inferred directly from other methods.

Scatter plots can be used for both qualitative and quantitative variables. The visual inspection of the scatter plots help to identify possible thresholds and non-linearity in response.

Scatter plots may be helpful in clarifying any interaction effect between inputs. For example, in Section 8.3.2 scatter plots were used to identify the interaction between the cooking

temperature and the precooking treatment in the cooking effect part. It is typically easier to make an inference regarding an interaction effect if at least one of the inputs involved in the interaction is qualitative.

Scatter plots may not be useful in developing quantitative rankings of the importance of different inputs. One reason for this limitation is that is difficult to evaluate the relative likelihood of outcomes in situations where it is possible that either few or many data points may overlap near the same coordinates of the graph. Thus, while one can make an inference regarding the general trend of the output with respect to the input, and regarding the range of values of the output that is not explained by changes in the input, it can be difficult to ascertain whether the ranges represent highly probable outcomes. Automatic pattern recognition algorithms may be used to aid in identifying patterns in the output that may or may not be easily discovered from direct visualization. However, these algorithms are not straight forward to implement. Commercial software such as *Partek* from pattern recognition and recognition technologies division of 'Partek Incorporated' and *PRTOOLS* from the pattern recognition toolbox in Matlab are available but their usefulness depends on the nature of actual dataset obtained. Furthermore, if an output responds in a qualitatively different manner to different inputs, such as because there may be a threshold effect with some inputs, nonlinear effects with other inputs, and so, it can be potentially difficult to reconcile these differences into an unambiguous ranking. However, judgments can be made and explained regarding which of these effects was of most concern to a particular decision maker or analyst. Thus, it is possible to use scatter plots as a basis for explaining judgments regarding which inputs are deemed to be of the most importance. Moreover, the inspection of scatter plots can help clarify results obtained from other methods. For example, inspection of scatter plots may help confirm that a breakpoint identified in CART actually represents a threshold.

Scatter plot were not used for two-dimensional analysis of variability and uncertainty. Plotting and interpretation of scatter plots is typically done manually, which makes it impractical for the large number of datasets involved in two-dimensional analysis.

### 20.1.8 Conditional Sensitivity Analysis Methods

Conditional sensitivity analysis is used to assess possible trends in the data and potentially complex dependencies between inputs and the outputs of interest. Conditional sensitivity analysis was implemented in order to clarify special relationships between the output

and inputs such as non-linearity in the model response and thresholds. Conditional sensitivity plots cannot explicitly be implemented in order to rank the inputs, because it can be difficult to assess the likelihood of the specific cases that are visualized. Furthermore, it can be difficult to compare responses to different inputs in situations in which the nature of the response is qualitatively different (e.g., a threshold effect in one case, a nonlinearity in another, a linear relationship in yet a third, and so on). However, the possibility of identifying and clarifying special relationships such as these is a key advantage of this method of sensitivity analysis. Therefore, this method can be a useful complement to other methods.

Conditional sensitivity analysis methods are typically not applied for quantitative inputs. A subset of the combination of inputs leading to high exposure risk cases can be identified using conditional sensitivity analysis. However, there may be high risk exposure cases that were not simulated during conditional sensitivity analysis, since such analyses are typically based upon a limited number of combinations of possible values. In most respects, the advantages and disadvantages of conditional sensitivity analysis are similar to those for scatter plots, with the key exception that conditional sensitivity analysis is intended to clarify interaction effects.

### 20.1.9 Summary of the Key Characteristics of Sensitivity Analysis Methods

Based on the discussion in Section 20.1.1 to 20.1.7, the key characteristics addressed by NRSA, DSA, sample and rank correlation, linear and rank regression analysis, ANOVA, CART analysis, scatter plots and conditional sensitivity analysis methods are summarized in Table 20-1. Twelve characteristics as listed in the table were evaluated based on this study.

Ideally, a sensitivity analysis method should respond to the effects of simultaneous variation in all inputs. The methods were evaluated to test 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 critical for food safety process risk models. Some methods handle only quantitative inputs and other can address both quantitative and qualitative inputs.

Because high exposure cases are of special interest, methods that can help identify and characterize conditions leading to high exposures are advantageous. The ability to address uncertainty provides robustness and confidence in the control measure applied based upon insights from the analysis. For example, the capability to perform sensitivity analysis easily in a

Table 20-1. Summary of Key Characteristics of Sensitivity Analysis Methods

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

two dimensional probabilistic simulation framework helps in assessing the key sources of variability and whether the identification of the most important inputs is robust with respect to uncertainty. 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 already exist, especially if they have user-friendly interfaces ('push button methods'). 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. However, 'push button' methods such correlation coefficients do not account for typical characteristics of food safety risk assessment models such as interaction and thresholds. Thus, even though regression, ANOVA and CART may be more difficult to apply than some of the readily available methods such as correlation coefficients, their use may be necessary to capture important characteristics of food safety risk assessment models.

The ability to produce quantitative rankings and, furthermore, the ability to evaluate the statistical significance of the rankings, is useful to the identification of the relative importance of inputs and the confidence that should be imputed to the rankings. Some methods produce more useful measures via 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 whether they produce reasonable results. Thus, the ability of the methods to produce reasonably correct results in practice, even if there were departures from key assumptions of the method, was assessed.

The two-dimensional analysis of variability under uncertainty provided a range of ranks for each of the inputs. The confidence in ranks in the face of uncertainty can be judged based upon the range of ranks and the probabilities associated with each of those ranks. Some inputs were consistently ranked high, whereas others were consistently related low. The information about range of ranks and associated probabilities is useful to a decision maker. For example, for the pâtés food category in *Listeria monocytogenes* model, storage time is always among the last two inputs and the probabilities o fit being valued last and second last are 0.66 and 0.33, respectively. Thus, the decision maker can safely conclude that it is not important to control storage time in this particular case.

Based upon these criteria and the judgments made regarding how well each method addresses each of the criteria, it is clear from Table 20-1 that there is no method that perfectly addresses all criteria, nor is there any criterion that is addressed by every one of the methods. Thus, there are trade-offs regarding the selection of sensitivity analysis methods.

It is clear that the mathematical methods of NRSA and DSA offer few theoretical advantages, with the exception of providing quantitative rankings and perhaps insight regarding distinctions among the rankings. These methods can be easy to apply depending upon the structure of the model itself. These methods were less reliable than others in providing reasonably correct ranking of the inputs.

The four statistical methods of sample and rank correlation, linear and rank regression analysis, ANOVA, and CART, were comparable in performance in many respects. However, ANOVA and CART are better suited to dealing with nonlinearity, thresholds, and interactions. Correlation coefficients, regression analysis and ANOVA provide a clear indication of the quantitative rank order of importance of inputs and regarding the robustness of the rankings.

All four of these methods provided reasonable results in practice based upon the case studies of this work. However, of these four methods, the theoretical basis of sample correlation and linear regression is the weakest with respect to application to nonlinear food safety process risk models. Furthermore, all four do not deal with thresholds. Thus, users should be cautious about the application of sample correlation and linear regression. Generally, higher the obtained value for $R^2$, more is the confidence in the conclusions based on the results. Rank regression and rank correlation theoretically would perform better for non-linear models but only in the case that the response is monotonic.

Correlation coefficients, regression analysis and ANOVA can be automated for two-dimensional analysis and hence provided probability of rank associated with a single sampling based on the number of test samplings. However, the reliability of results based on ANOVA in the two-dimensional case must be tested based on the discussion in Section 20.1.5 pertaining to definition of levels. Although correlation coefficients and regression analysis appear to address similar characteristics in the table, correlation coefficients do not address interactions and thresholds. For regression analysis not restricted by a linear model, terms can be added to

account for interactions and change point regression can be used to handle thresholds. Thus, the reliability of ranking obtained from correlation analysis is less than that of regression analysis.

The statistical methods of regression analysis, ANOVA and CART vary in complexity and run time. Regression analysis typically involves least amount of complexity as it involves standardization of data and then solution of 'n' equations in terms of 'x' variables, where 'n' is the number of simulations and 'x' the number of inputs. ANOVA on the other hand has similar complexity but takes more amount of time to allow calculation of contrasts. However, since the numbers in ANOVA are categorical in nature, the calculation of F-values for the factors is relatively fast. CART involves calculation of deviance but there is no matrix inversion to get the estimates for sensitivity index. But in order to identify correct splits and convergence towards a homogeneous dataset, numerical search procedures are used. These search procedures have higher complexity compared to matrix inversion.

## 20.2 Recommendations

This study identified several key considerations regarding the development of food safety process risk models and the selection of sensitivity analysis methods. With regard to the former, in order to facilitate sensitivity analysis, food safety process risk models should be structured and programmed so as to clearly distinguish between inputs, the mathematical model, and outputs. Models implemented in programming languages such as Visual Basic are preferred over those implemented in spreadsheet environments as they allow for easier inspection, understanding and modification of the code. In particular, it is easier to interface the food safety process risk model with sensitivity analysis tools if the model adheres to good practices of software design and documentation.

Two dimensional analyses of variability and uncertainty are typically preferred over one dimensional approaches because of their theoretical appeal in correctly distinguishing between variability and uncertainty and in more appropriately enabling assessment of risk management issues. Furthermore, as demonstrated by case studies in this work, the estimation of sensitivity in a two-dimensional simulation provides key insights regarding the robustness of results and, hence, the confidence that can be placed in the rankings. This insight is needed to support decision making under uncertainty. For example, the range of ranks and probability distribution of ranks of an input is of valuable importance to the decision maker. However, sensitivity

analysis in a two dimensional simulation is potentially computationally intensive. Furthermore, as a practical manner, it is difficult to apply sensitivity analysis methods in a two dimensional framework unless they can be automated.

This study demonstrated sensitivity analysis in a two dimensional framework. The automation was done using SAS© macros. Application of sensitivity analysis to one-dimensional simulations took less time compared to two-dimensional analysis and did not require the same degree of automation. A comparison of one-dimensional and two-dimensional results implied that the top ranked input was the same for each of the probabilistic simulation approaches. Thus, it may be possible to develop a robust insight regarding key inputs using a simplified probabilistic simulation approach if resources do not permit a two dimensional approach. Alternatively, as part of model development, sensitivity analysis can initially be performed in a one-dimensional analysis to identify priorities for data collection. However, the results obtained in the specific case studies of this work may be indicative but are not definitive. Thus, for two dimensional models, a two dimensional sensitivity analysis is generally preferred unless a one-dimensional analysis is appropriately justified.

The statistical-based methods were found to be the most appropriate for application to food safety risk assessment models. In particular, ANOVA and CART are best able to deal with simultaneous variation in all inputs, qualitative and quantitative inputs, nonlinearities, interactions, and thresholds. ANOVA provides a quantitative ranking of individual factors in a multi-factor analysis. However, it is more complicated to compare the importance of interactions with regard to the importance of individual factors. With CART, the inputs that are selected in the tree can be ranked using quantitative measures and those, which are not, are less important than those in the tree. However, the relative importance of those that are not selected cannot be ranked. Furthermore, it is possible to identify combinations of values of inputs that produce the largest (or smallest) mean response. Thus, it is possible to prioritize situations that lead to the worst outcomes and, therefore, to develop insight as to how to avoid such outcomes by controlling the values of key inputs.

While often overlooked, it is also important to identify inputs that are not important to an assessment. This is because determining which inputs do not influence the results can save a significant amount of time as no additional information or strategy development is necessary for

such inputs. Thus, the identification of unimportant inputs is useful to the prioritization of resources toward other inputs that actually can make a difference in risk management. The statistical-based methods are most useful in identifying unimportant inputs with an acceptable degree of confidence, especially in the case of ANOVA.

The form of regression analysis most typically used in this work was linear standardized sample regression. While this approach can deal with simultaneous variation in all inputs, it is more challenging to interpret the results for qualitative inputs versus quantitative ones. An approach for ranking key inputs based upon F values was evaluated. This approach enables comparison of quantitative and qualitative inputs. The approach performed reasonably well. However, because the regression approach employed here imposes a linearity assumption, the fit of the regression model to the food safety process risk model was poor in some cases. A poor fit implies that the results of the analysis may be incorrect. Nonetheless, the results from regression analysis agreed well with the results of CART and ANOVA in most cases. Thus, in practice, the use of linear standardized sample regression appeared to be robust to departures from linearity. However, these results could be different for other food safety process risk models aside from those evaluated in this study. Other methods of regression analysis, such as rank regression or regression using nonlinear basis functions, could be used to obtain improved coefficients of determination and, therefore, to provide potentially more robust insights regarding sensitivity.

The mathematical methods of NRSA and DSA are based upon a linearity assumption. More importantly, these methods do not account for simultaneous variation in all inputs. The results from these two methods differed from those of the statistical methods in several cases. Furthermore, DSA does not provide a measure of the importance of each input. It merely indicates how substantially the model responds to an arbitrary small perturbation in an input, which we refer to as an indication of sensitivity rather than importance. It can be useful to compare the results of DSA with other methods to determine whether it is the model structure or perhaps a wide range of values for a given input that is responsible for an input being ranked as highly important by other methods.

The use of graphical methods, such as conditional sensitivity analysis, scatter plots, or both, is highly recommended. These techniques have the key advantage of providing insight regarding complex interactions between inputs, and regarding the response of the output to an

input, that cannot easily be inferred by the other methods. Thus, although the graphical methods do not provide a clear ranking of inputs, they provide information that is useful in interpreting the results of other methods.

Sensitivity analysis methods such as FAST, MII and Sobol's indices were not applied in this study for reasons mentioned in Chapter 2. They are theoretically promising methods and may merit consideration once the required software and resources are available. Other useful of sensitivity analysis methods may emerge that would merit evaluation.

This study demonstrated the capability of different sensitivity analysis methods to identify the relative importance of potential critical control points. However, from a management perspective not all potential CCP's can easily be addressed. If effort is put into modeling things that can directly affect risk management (RM) decisions then the insights drawn from the application of sensitivity analysis will be of direct decision importance. For example, based upon the risk assessment and sensitivity analysis, the risk manager should gain perspective as to what level of sanitation control should be applied in the slaughterhouse or as to what sort of public awareness/educational program should be formulated. It is also useful to know what part of the risk is beyond direct control. However, risk management problems are often complex and stakeholders can have multiple agendas. Moreover, the users of a particular assessment are not necessarily going to stick to the original assessment objectives. Therefore, it is useful to characterize the sensitivity of the model to uncertainty or variation in assumptions that might be subject to change by different users with different assessment objectives. Knowledge of key sources of uncertainty in the model can be used to prioritize research in order to reduce uncertainty. Thus, risk assessment, aided by sensitivity analysis, should include both a decision making and a research planning relevance.

**REFERENCES**


Alavi, S. H., V. M. Puri, S. J. Knabel, R. H. Mohtar, and R. C. Whiting (1999), "Development and validation of a dynamic growth model for *Listeria monocytogenes* in fluid whole milk," *Journal of Food Protection* 62(2): 170-176.

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.

Ariens, G.A., W. van Mechelen P.M. Bongers, L.M. Bouter, and G. van der Wal (2000), "Physical Risk Factors for Neck Pain," *Scandinavian Journal of Work Environment & Health*, 26(1): 7-19.

Bischof, C.H., A. Carle, G. Corliss, A. Griewank, and P. Hovland (1992). "ADIFOR Generating Derivative Codes From FORTRAN Programs." ANL-4501, Prepared by Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Illinois.

Borkman, M., Storlien, L.H. et al. (1993), "The relation between insulin sensitivity and the fatty-acid composition of skeletal-muscle phospholipids," *The New England Journal of Medicine*, 328(4): 238-44.

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

Brikes, D., and Y. Dodge (1993). *Alternative Methods of Regression*. John Wiley and Sons: New York.

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.

Buchholz, U. (2000). Los Angeles birth and population data from 1985-1999. Data includes perinatal listeriosis deaths per year and listeriosis deaths per year (unpublished data).

Campolongo F. and A. Rossi (2002), "Sensitivity Analysis and the Delta Hedging Problem," Proceedings of the Second Summer School on Sensitivity Analysis of Model Output (SAMO), Venice. Sponsored by the European Safety and Reliability Association and the IEEE Reliability Society - Italian Chapter.

Carmichael, G.R., A. Sandu, and F.A. Potra (1997), "Sensitivity Analysis for Atmospheric Chemistry Models Via Automatic Differentiation," *Atmospheric Environment*, 31(3):475-489.

Centers for Disease Control and Prevention (2000). Preliminary FoodNet Data on the Incidence of Foodborne Illnesses – Selected Sites, United States, 1999. *Morbidity and Mortality Weekly Report*, 49(10): 201-205.

Centers for Disease Control and Prevention (1998a). The 1997 Annual Report by Center for Diseaese Control and Prevention/United States Department of Agriculture/Food and Drugs Administration, Foodborne Diseases Active Surveillance Network CDC's Emerging Infections Program. http://www.cdc.gov/foodnet/ANNUAL/97_surv.htm.

Centers for Disease Control and Prevention. 1998b. Multistate outbreak of listeriosis - United States, 1998. *Morbidity and Mortality Weekly Report* 47(50):1085-1086.

CFSAN (2001b), Relative Risk to Public Health from Foodborne *Listeria monocytogenes* Among Selected Categories of Ready-to-Eat Foods, Draft Risk Assessment Document and Risk Management Action Plan, Prepared by Center for Food Safety and Nutrition/Food and Drugs Administration/ U. S. Department of Agriculture, Availability Federal Register Notice. January 19, 2001. http://www.fsis.usda.gov/OPHS/lmrisk/DraftLm22603.pdf, Viewed August 10, 2003.

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.

Chen, P., P. M. Popovich (2002). *Correlation Parametric and Nonparametric Measures*, Sage Publications, Inc.: Thousands Oaks, CA.

Collins, D.C., and R. Avissar (1994), "An Evaluation with the Fourier Amplitude Sensitivity Test (FAST) of Which Land-Surface Parameters Are of Greatest Importance in Atmospheric Modeling," *Journal of Climate*, 7(5): 681-703.

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.

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.

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

Duh, Y.H. and D. W. Schaffner (1993), "Modeling the effect of temperature on the growth rate and lag time of *Listeria innocua* and *Listeria monocytogenes*," *Journal of Food Protection* 56(3): 205-210.

DHHS (1997). *National Health and Nutrition Examination Survey, III 1988-94*, version CD-ROM Series 11, No. 1A.Centers for Disease Control and Prevention, U.S. Department of Health and Human Services, National Center for Health Statistics Hyattsville, MD.

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.

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.

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.

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., 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. (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., 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

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

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.

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

Grievank, A., (2000), *Principles and Techniques of Algorithmic Differentiation*: *Evaluating Derivatives*, SIAM Publisher: Philadelphia.

Hays, J. F., D. Walker, B.H. Hager, E. M. Stopler (1981), "Melt segregation from partially molten source regions: The importance of melt density and source region size," *Journal of Geophysical Research*, 86:6261-6271.

Helton, J.C. (1993), "Uncertainty and Sensitivity Analysis Techniques for Use in Performance Assessment for Radioactive-Waste Disposal," *Reliability Engineering and System Safety*, 42(2-3): 327-367.

Helton, J.C., and R.J. Breeding (1993), "Evaluation of Reactor Accident Safety Goals," *Reliability Engineering & System Safety*, 39(2): 129-158.

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.

Hora, S. C., and R. L. Iman (1990), "A Robust Measure of Uncertainty Importance for Use in Fault Tree System Analysis," *Risk Analysis*, 10(3): 401-406.

Hotelling H. and M.R. Pabst (1936). "Rank Correlation and Tests of Significance Involving No Assumption of Normality," *Annals of Mathematical Statistics*, 7, 29-43.

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

Hwang, D., D.W. Byun, and M.T. Odman (1997), "An Automatic Differentiation Technique for Sensitivity Analysis of Numerical Advection Schemes in Air Quality Models," *Atmospheric Environment*, 31(6): 879-888.

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.

Kendall, M. G. and G. U. Yule (1948). *An Introduction to the Theory of Statistics*. Charles Griffin and Company. London.

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

Kendall, M.G., and J.D. Gibbons (1990). *Rank Correlation Methods*. Fifth Edition. Charles Griffin: London.

Kim, J.O., and C.W. Mueller (1978). *Introduction to Factor Analysis: What It Is and How to Do It*. Sage Publications: Beverly Hills, CA.

Kohn, M.C. (2000), "Use of Sensitivity Analysis to Assess Reliability of Metabolic and Physiological Models," *Risk Analysis*, 22 (3): 623-631.

Krishnaiah, P.R (1981). *Ananlysis of Variance*. Elsevier: New York.

Krzykacz, B. (1990). "SAMOS: A Computer Program for the Derivation of Empirical Sensitivity Measures of Results from Large Computer Models." GRS-A-1700, Gesellschaft Fur Reaktorsicherheit (GRS) mbH, Garching, Germany.

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

Lindman, H.R (1974). *Analysis of Variance in Complex Experimental Designs*. W. H. Freeman & Co.: San Francisco, CA.

Liu, Y.Q., and R. Avissar (1996), "Sensitivity of Shallow Convective Precipitation Induced by Land Surface Heterogeneities to Dynamical and Cloud Microphysical Parameters," *Journal of Geophysical Research-Atmospheres*, 101(D3): 7477-7497.

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.

McNemar, Q. (1969), Psychological Statistics (4th ed.), New York Wiley.

McRae, G.J., J.W. Tilden, and J.H. Seinfeld (1982), "Global Sensitivity Analysis – A Computational Implementation of the Fourier Amplitude Sensitivity Test," *Computers and Chemical Engineering*, 6(1): 15-25.

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.

Mohanty, S., and T.J. McCartin (1998). *Total-System Performance Assessment (TPA) Version 3.2 Code: Module Descriptions and User's Guide.* Center for Nuclear Waste Regulatory Analyses: Southwest Research Institute, Texas.

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.

Montgomery, D.C. (1997). *Design and Analysis of Experiments*. Wiley and Sons Ltd.: New York.

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

Ogden, T., E. Parzen (1996). "Data Dependent Wavelet Thresholding in Nonparametric Regression with Change-point Applications," *Computational Statistics and Data Analysis* 22(1): 53-70.

Olds, E. G. (1949). The 5 percent significance levels for sums of squares or ranked differences and a correction. *Annals of Mathematical Statistics*, 20, 117-118.

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).

Pate-Cornell, E. (2002), "Risk and Uncertainty Analysis in Government Safety Decisions," *Risk Analysis*, Vol. 22. No. 3.

Patil, S.R., and H.C. Frey (2003), "Comparison of Sensitivity Analysis Methods Based Upon Applications to a Food Safety Risk Model," *Risk Analysis*, submitted and undergoing peer review.

Ratto M., N. Giglioli, S. Tarantola, U. Callies, Å. Forsman (2001a), "Sensitivity Analysis in Calibration and Reduction of Environmental Models: Eutrophication Case Studies," ESREL 2001, Safety and Reliability, Proceedings of the European Conference on Safety and Reliability E. Zio, M. Demichela, N. Piccinini (eds.), Torino, vol. 1, 301-308.

Ratto, M., S. Tarantola, A. Saltelli (2001b), "Sensitivity Analysis in Model Calibration: GSA-GLUE Approach," *Computer Physics Communication*, 136(3): 212-224.

Ratkowsky, D. A., J. Olley, T. A. McMeekin, and A. Ball. (1982), "Relationship between temperature and growth rate of bacterial cultures," *Journal of Bacteriology* 149:1-5.

Riley, Remis, Helgerson (1983), "Hemorrhagic colitis associated with a rare Escherichia coli serotype," *N Engl J Med*, Mar 24, 308(12):681-685.

Rodgers, J. L., and Nicewander, W. A. (1988), "Thirteen Ways to Look at the Correlation Coefficient," *The American Statistician*, 42(1):59-66.

Rodriguez-Camino, E., and R. Avissar (1998), "Comparison of Three Land-Surface Schemes with the Fourier Amplitude Sensitivity Test (FAST)," *Tellus Series A - Dynamic Meteorology and Oceanography*, 50(3): 313-332.

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

Salehi, F., S.O. Prasher, S. Amin, A. Madani, S.J. Jebelli, H.S. Ramaswamy, and C. T. Drury (2000), " Prediction of Annual Nitrate-N Losses in Drain Outflows with Artificial Neural Networks," *Transactions of the ASAE*, 43(5): 1137-1143.

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. (2001), "Sensitivity Analysis for Importance Assessment," Proceedings of North Carolina State University/United States Department of Agriculture Workshop on Sensitivity Analysis Methods, Raleigh, North Carolina. www.ce.ncsu.edu/risk.

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.

Siegel, S. (1956) *Nonparametric Statistics for the Behavioural Sciences*, McGrawHill: New York.

SimLab (2000) Simluation Environment for Uncertainty and Sensitivity Analysis, Version 1.1. Developed by the Joint Research Centre, Ispra, Italy. http://www.jrc.cec.eu.int/uasa/prj-sa-soft.asp. 07-09-2003.

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.

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.

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.

Thompson, K.M. (2002), "Variability and Uncertainty Meet Risk Management and Risk Communication," Risk Analysis, 22(3): 647-654

USDA (1997). *Pathogen Modeling Program*. Agricultural Research Service, United States Department of Agriculture, Wyndmoor, PA. www.arserrc.gov/mfs/.

USDA (1998a). *Design and Operation: The Continuing Survey of Food Intakes by Individuals and the Diet and Health Knowledge Survey, 1994-1996*: Nationwide Food Surveys Report, Agriculture Research Service, U.S. Department of Agriculture. Springfield, VA: National Technical Information Service.

USDA (1998b). *1994-96 Continuing Survey of Food Intakes by Individuals and 1994-96 Diet and Health Knowledge Survey and Technical Support Databases.* National Technical Information Service, Agriculture Research Service, United States Department of Agriculture. Springfield, VA.

Vose, D. J., (1998). "The Application of Quantitative Risk Assessment to Microbial Food Safety," *Journal of Food Protection* 61(5):640-648.

Walpole, R. E. and R. H. Myers (1993). Probability and Statistics for Engineers and Scientists, New York: MacMillan Publishing Company.

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.

Winter, B.J., D.R. Brown, and K.M. Michels (1991). *Statistical Principles in Experimental Design*. Third Edition. McGraw-Hill Inc.: New York.

Wijtzes, T., P. J. McClure, M. H. Zwietering, and T. A. Roberts (1993), "Modelling bacterial growth of *Listeria monocytogenes* as a function of water activity, pH and temperature," *International Journal of Food Microbiology*, 18:139-149.

WHO (1988), "Foodborne Listeriosis." World Health Organization Working Group, *Bulletin of the World Health Organization,* 66: 421-428.