IDENTIFICATION AND REVIEW OF SENSITIVITY ANALYSIS METHODS

By

H. Christopher Frey, Ph.D.

Civil Engineering Department

North Carolina State University

Raleigh, NC.

Sumeet R. Patil

Civil Engineering Department

North Carolina State University

Raleigh, NC.

Abbreviated Title: Sensitivity Analysis Methods

Correspondence:

H. Christopher Frey

Civil Engineering Department

311 Mann Hall, Campus Box 7908

North Carolina State University

Raleigh, NC 27695-7908.

Email: frey@eos.ncsu.edu

ABSTRACT

Identification and qualitative comparison of sensitivity analysis methods that have been used across various disciplines, and that merit consideration for application to food safety risk assessment models, are presented in this paper. Sensitivity analysis can help in identifying critical control points, prioritizing additional data collection or research, and verifying and validating a model. Ten sensitivity analysis methods, including four mathematical methods, five statistical methods and one graphical method, are identified. The selected methods are compared on the basis of their applicability to different types of models, computational issues such as initial data requirement and complexity of their application, representation of the sensitivity, and the specific uses of these methods. Applications of these methods are illustrated with examples from various fields. No one method is clearly best for food safety risk models. In general, use of two or more methods, preferably with dissimilar theoretical foundations, may be needed to increase confidence in the ranking of key inputs.

Key Words: Sensitivity analysis methods, food safety, microbial risk assessment, critical control points.

1.0 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, and an increasing global food trade (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 5,200 deaths. ERS (2001) estimated the cost of food borne disease to be $6.9 billion annually. Food safety is gaining increased attention because of several trends. New food borne pathogens are emerging (Buchanan, 1996). Larger batch production, distribution, and longer shelflife of food products contribute to broader exposure to contaminating events. The proportion of food consumed that is supplied by food services is growing and this reduces consumer control over food handling and processing. Demand for safer food is growing as consumers are becoming more affluent, live longer, and are better informed about diet (McKone, 1996). International trade in food also can introduce new sources of risks to food safety, such as cattle or meat acquiring infection overseas.

Food safety regulatory agencies are taking a new approach to ensuring the safety of food supply based upon the Hazard Analysis Critical Control Points (HACCP) system (Seward, 2000). One step in the HACCP system is to determine critical control points (CCP) where risk management efforts can be focused. Food safety depends on many factors, such as composition and preparation of the product, process hygiene, and storage and distribution conditions (Zwietering and Gerwen, 2000). Given data gaps and the complexity and dynamic nature of the food processing, transportation, storage, distribution, and preparation system, determining which of the many nodes in the farm-to-table pathway constitute CCPs and represents 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 and aid in developing priorities for risk mitigation. For example, Baker *et al.* (1999) identified sensitivity analysis as one of the principal quantitative techniques used for risk management in the United Kingdom. Jones (2000) indicated that sensitivity analysis can provide the basis for planning adaptation measures to mitigate the risk of climate change. Sensitivity analysis can be used as an aid in identifying the important uncertainties for the purpose of prioritizing additional data collection or research (Cullen and Frey, 1999). Furthermore, sensitivity analysis can play an important role in model verification and validation throughout the course of model development and refinement (*e.g*., Kleijnen, 1995; Kleijnen and Sargent, 2000; and Fraedrich and Goldberg, 2000). Sensitivity analysis also can be used to provide insight into the robustness of model results when making decisions (*e.g.*, Phillips *et al.*, 2000; Ward and Carpenter 1996; Limat *et al.*, 2000; Manheim 1998; and Saltelli *et al*., 2000). Sensitivity analysis methods have been applied in various fields including complex engineering systems, economics, physics, social sciences, medical decision making, and others (*e.g.*, Oh and Yang, 2000; Baniotopoulos, 1991; Helton and Breeding, 1993; Cheng, 1991; Beck *et al.*, 1997; Agro *et al*., 1997; Kewley *et al*., 2000; Merz *et al*., 1992).

1.1 Objectives and Motivation.

The objectives of this paper are to identify, review, and evaluate sensitivity analysis methods for applicability to risk assessment models typical of those used, or expected to be used, in the near future for food safety risk assessment. As the key objective in this task is to benefit from knowledge across many disciplines, the literature review extends well beyond food safety to include many other fields. This objective serves as an aid in identifying potential CCPs 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.

1.2 Importance of Risk Assessment in Food Safety.

The objective of the risk assessment is to answer three risk questions (Kaplan and Garrick, 1981):

- What can go wrong?
- How likely is that to happen?
- What would be the consequences if it did go wrong?

Risk assessment provides structured information that allows decision makers to identify interventions that can lead to public health improvement or to avoid future problems.

The HACCP system is used to prevent hazards associated with foods. The best information of the risk from such hazards that industry may have today is qualitative, such as whether a hazard presents a high, medium, or low risk (WHO, 2001). The outcome of a HACCP-based system should be improved food safety assurance, measured by reduction in risk to the consumer and not merely a reduction in the level of a hazard in a food (Hathaway, 1995). Risk assessment can be used to determine which hazards are essential to control, to reduce, or to eliminate (Buchanan, 1995; Hathaway, 1995; and Notermans *et al.*, 1995). Therefore, risk assessment can help in developing more effective HACCP plans.

Risk assessments can play an important role in international trade by ensuring that countries establish food safety requirements that are scientifically sound and by providing a means for determining whether different standards provide equivalent levels of public health protection (WHO, 2001). Without a systematic risk assessment, countries may set requirements that are not related to food safety and can create artificial barriers to trade.

1.3 Risk Assessment Framework.

As defined by the Codex Alimentarius Commission (CAC), risk assessment is a scientifically based process consisting of four main steps: hazard identification, hazard characterization, exposure assessment, and risk characterization (FAO, 1999). These steps are similar to those defined in the risk analysis framework of the National Academy of Science (NRC, 1983). Each of these four steps is described briefly according to CAC definitions.

- Hazard Identification. Hazard identification involves identification of biological, chemical, and physical agents capable of causing adverse health effects and that may be present in a particular food or group of foods.

- 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 Assessment. Exposure assessment is the qualitative and/or quantitative evaluation of the likely intake of biological, chemical, and physical agents via food as well as exposures via other sources, if relevant.

- Risk Characterization. Risk characterization involves qualitative and/or quantitative estimation, including attendant uncertainties, of the probability of occurrence and severity of known or potential adverse health effects in a given population based on hazard identification, hazard characterization, and exposure assessment.

1.4 Important Issues in Food Safety Modeling.

There are many important issues in modeling as applied to risk assessment that motivate the need for sensitivity analysis and that illustrate the key challenges that sensitivity analysis faces or that can be addressed by sensitivity analysis. Cullen and Frey (1999) discuss issues such as purpose of models, model extrapolation, inappropriate application of models, complex models, and model verification and validation.

1.4.1 Purpose of the Model.

The purpose of a model is to represent as accurately as necessary a system of interest. A system is typically comprised of many components. All modeling involves decisions regarding aggregation and exclusion. Aggregation refers to simplified representation of complex real world systems. Exclusion refers to a decision to omit any portions that are judged not to be important with respect to modeling objectives (Cullen and Frey, 1999).

The reason that risk models are developed is to serve as an aid in decision making. Models are developed for different purposes, often with different decision making objectives in mind. For example, three different objectives are addressed by screening analysis models, research models, and assessment/decision making models, respectively.

Screening analysis is usually based on simple models that are not likely to underestimate the risk. The purpose of this analysis is usually to help the decision maker with routine regulatory decisions. Screening methods are intended to help identify exposure pathways that are not important and to do so with a great deal of confidence. Screening methods cannot be used to prove that a particular exposure pathway is important. Rather, they can identify exposure pathways that should receive priority for further evaluation and analysis. Because screening models are conservative, they are designed to provide some "false positives." This means that they will often provide results indicating that a exposure pathway is of concern, but later analysis with more refined assumptions or models may reveal less of a problem. A key advantage of screening models is that they are often easier to use than more refined models. Screening models typically have fewer input data requirements and involve less complex calculations.

Research models are intended to improve understanding of the function and structure of real systems. They allow a researcher to explore possible and plausible functional relationships and may involve many detailed mechanisms. Research models may be complicated and may not be aimed at any particular risk management end point. As such, they may have large input data requirements, be difficult to execute, and provide output not directly relevant to a specific risk management objective. However, such models can be very helpful in improving fundamental insights regarding risk processes and in learning lessons useful for further model development.

Refined assessment models are intended to serve as tools for decision making, such as for rule-making or regulatory compliance purposes where screening analyses are inadequate. These models are generally complex and more accurate than screening models. Refined models are developed with at least some attention to risk management application. Compared to screening models, refined models typically have greater data requirements and require more time and experience to apply.

Sensitivity analysis is important for all three types of models. In the context of this discussion, refined risk models are perhaps the most relevant. 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.

1.4.2 Complex Models.

A model can be small yet complex. Model size should not be confused with model complexity. Complex systems are often hierarchies, which can be described in terms of the "span" of each level in the hierarchy and the number of levels (Reed and Afjeh, 2000; Simon, 1996).

Food Safety Risk Models (FSRMs) typically: (1) are nonlinear; (2) contain discrete inputs; (3) contain thresholds; (4) contain multiple pathways; and (5) are modular. The nonlinear and threshold features imply that interactions are important. For example, if temperature is low enough, then there may be no microbial growth, but once temperature exceeds a threshold, then the model may predict a nonlinear response that depends on several model inputs. The modularity feature means that computations may take place in separate modules of the models, and only a selected set of aggregated results may be passed from one module to another.

1.4.3 Model Verification and Validation.

Model verification is a process of making sure that the model is doing what it is intended to do. Sensitivity analysis can be helpful in verification. If a model responds in an unacceptable way to changes in one or more inputs, then troubleshooting efforts can be focused to identify the source of the problem.

Model validation ideally involves comparison of model results to independent observations from the system being modeled. Generally, in most risk assessments, complete validation is not possible because of lack of sufficient observational data. Risk assessment models usually predict rare events. For example, in radiation exposure assessments, model validation may be nearly impossible due to extremely low levels of radionuclide concentrations in the environment or because the time periods considered by the model are prohibitively long (Hoffman and Miller, 1983). Cullen and Frey (1999) discuss partial validation of a model when observational data are available for only a part of the modeling domain.

Sensitivity analysis can be used to help develop a "comfort level" with a particular model. If the model response is reasonable from an intuitive or theoretical perspective, then the model users may have some comfort with the qualitative behavior of the model even if the quantitative precision or accuracy is unknown.

1.4.4 Model Extrapolation.

A model is applicable within the specific set of inputs and outputs associated with the data used for calibration or validation of the model. Model extrapolation involves making predictions for a situation beyond the range of calibration or validation of the model. For example, Guassian-plume-based air dispersion models are applicable only for distances up to about 20 km or 50 km from the emission source. The model should not be used to characterize long-range transport over hundreds of kilometers.

Empirical models such as ones developed from regression analysis are generally based on a specific data set. Regression analysis may be based on arbitrary functions to approximate relationships between the data. However, there may not be any theoretical basis for the functional relationship selected. In such a situation, if the inputs used are outside the original data set from which the regression model was developed then the output may not valid. There are two types of extrapolations. Explicit extrapolation involves making predictions for values of inputs outside the range of values for which the model was calibrated or validated. Hidden extrapolation involves specifying combinations of inputs for which validation has not been done, even though each of the input values fall within the range of values that have been tested. If the functional form of the relationship between the output and the inputs of the model is based on sound theoretical assumptions, then the model may perform reasonably well when extrapolated. Sensitivity analysis can be used to reveal how the model performs when the model is extrapolated.

1.5 Sensitivity Analysis Methods.

Sensitivity analysis methods can be classified in a variety of ways. In this report, they are classified as: (1) mathematical; (2) statistical; and (3) graphical. Other classifications focus on the capability, rather than the methodology, of a specific technique (*e.g.*, Saltelli *et al.*, 2000). Classification schemes aid in understanding the applicability of a specific method to a particular model and analysis objective. Here, the focus is on sensitivity analysis techniques applied in addition to the fundamental modeling technique. For example, an analyst may perform a deterministic analysis, in which case a mathematical method, such as nominal range sensitivity can be employed to evaluate sensitivity. Alternatively, an analyst may perform a probabilistic analysis, using either frequentist or Bayesian frameworks, in which case statistical-based sensitivity analysis methods can be used (*e.g*., Cullen and Frey, 1999; Box and Tiao, 1992; Saltelli *et al.*, 2000; Weiss, 1996).

1.5.1 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 that represent the possible range of the input (*e.g.*, Salehi *et al.*, 2000). These methods do not address the variance in the output due to the variance in the inputs, but they can assess the impact of range of variation in the input values on the output (Morgan and Henrion, 1990). In some cases, mathematical methods can be helpful in screening the most important inputs (*e.g.*, Brun *et al.*, 2001). These methods also can be used for verification and validation (*e.g.*, Wotawa *et al.*, 1997) and to identify inputs that require further data acquisition or research (*e.g.*, Ariens *et al.*, 2000). Mathematical methods evaluated here include nominal range sensitivity analysis, break-even analysis, difference in log odds ratio, and automatic differentiation.

1.5.2 Statistical Methods for Sensitivity Analysis.

Statistical methods involve running simulations in which inputs are assigned probability distributions and assessing the effect of variance in inputs on the output distribution (*e.g.*, Andersson *et al.*, 2000; Neter *et al.*, 1996). Depending upon the method, one or more inputs are varied at a time. Statistical methods allow one to identify the effect of interactions among multiple inputs.

The range and relative likelihood of inputs can be propagated using a variety of techniques such as Monte Carlo simulation, Latin hypercube sampling, and other methods. Sensitivity of the model results to individual inputs or groups of inputs can be evaluated by variety of techniques (Cullen and Frey, 1999). Greene and Ernhart (1993), Fontaine and Jacomino (1997), and Andersson *et al.* (2000) give examples of the application of statistical methods. The statistical methods evaluated here include regression analysis, analysis of variance, response surface methods, Fourier amplitude sensitivity test, and mutual information index.

1.5.3 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 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 representation (*e.g.*, Stiber *et al.*, 1999; Critchfield and Willard, 1986).

1.6 Organization of the Report.

Section 2 identifies, describes, and evaluates ten selected sensitivity analysis methods. Section 3 presents a comparison the methods. Conclusions are given in Section 4.

2.0 SENSITIVITY ANALYSIS METHODS.

This section identifies sensitivity analysis methods used across various disciplines. Reference materials, such as journals, reports, books, and the World Wide Web were collected from various fields including medical decision making, risk analysis, environmental science, engineering, economics, ecology, statistics, health science, microbiology and food science. Methods that were applied on models comparable to risk assessment models were identified for evaluation. The methods and specific applications of each method are demonstrated. Strengths and limitations of the methods are noted in brief.

2.1 Nominal Range Sensitivity.

Nominal range sensitivity method 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. It is usually not used for probabilistic analysis. One use of nominal sensitivity analysis is as a screening analysis to identify the most important inputs to propagate through a model in a probabilistic framework (Cullen and Frey, 1999). Nominal range sensitivity can be used to prioritize data collection needs as demonstrated by Salehi *et al.* (2000).

2.1.1 Description.

Nominal range sensitivity analysis evaluates the effect on model outputs exerted by individually varying only one of the model inputs across its entire range of plausible values, while holding all other inputs at their nominal or base-case values (Cullen and Frey, 1999). The difference in the model output due to the change in the input 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 results of nominal range sensitivity 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 nominal range sensitivity are potentially misleading for nonlinear models.

2.1.2 Application.

Two examples of the use of nominal range sensitivity analysis are given here. In the first, nominal range sensitivity was applied as described here to an exposure assessment problem for flounder in a contaminated harbor (Dakins *et al*., 1994). The selection of inputs evaluated in the sensitivity analysis, and the range of values assigned to those inputs, was based upon literature review. The absolute different in burden of PCB in flounder was the key measure of sensitivity. An absolute difference of less than 0.5 μg PCB per gram of body weight was considered to be non-sensitive.

In the second example, a conditional nominal range sensitivity analysis was performed (Sinicio *et al*., 1997). The objective of the analysis was to determine the allowable safe storage time for grain wheat as a function of grain moisture content and temperature conditioned upon key assumptions of a ventilation system. The intent was to recommend airflow rates and fan control strategies to properly aerate the stored grains. The sensitivity analysis was conditioned on two different assumptions regarding the airflow rate and two different assumptions regarding the air temperature rise in the fan, leading to four different cases. For each case, three specific inputs were each varied over a nominal range, and the impact on the key model output, allowable storage time elapsed, was evaluated. Specific heat of the wheat was observed to be the most sensitive input for each of the four cases.

2.1.3 Advantages.

Nominal range sensitivity analysis 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.

2.1.4 Disadvantages.

Nominal sensitivity analysis 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 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 small changes in a few or all inputs together are not shown by nominal sensitivity analysis For other than linear models, it is not clear that nominal range sensitivity can provide a reliable rank ordering of key inputs.

2.2 Difference in Log-Odds Ratio (ΔLOR).

The difference in the log odds ratio (ΔLOR) method is a specific application of nominal range sensitivity methodology. The ΔLOR is used when the output is a probability. For example, Song *et al.* (2000) used ΔLOR for identifying the most important inputs so that additional data collection and research can be prioritized.

2.2.1 Description.

The odds or odds ratio of an event is a ratio of the probability that the event occurs to the probability that the event does not occur (Walpole and Myers, 1993). If an event has a probability of occurrence as P, then the odds ratio is $P/(1-P)$. The log of the odds ratio or logit is just another convenient way of rescaling probabilities (Menard, 1995). Log odds are considered by some to be the preferred transformation of probability because they are putatively easier to understand (Christensen, 1990).

The ΔLOR method is used to examine the change in the output as:

$$\Delta LOR = \log\left[\frac{\Pr(event\,|\,with\ changes\ in\ input)}{\Pr(No\ event\,|\,with\ changes\ in\ input)}\right] \; - \; \log\left[\frac{\Pr(event\,|\,without\ changes)}{\Pr(No\ event\,|\,without\ changes)}\right].$$

If ΔLOR is positive, changes in one or more inputs enhance the probability of the specified event. If ΔLOR is negative, then the changes in the inputs cause a reduction in the probability of the event occurring or increase the probability of the event not occurring. The greater the magnitude of ΔLOR, the greater is the influence of the input (Stiber *et al.*, 1999).

2.2.2 Application

The *DLOR* method was used by Stiber *et al*. (1999) to identify key inputs to a model of groundwater decontamination via reductive dechlorination. The model inputs referred to as "evidences", included site parameters such as temperature, pH, and whether various specific chemicals were found to be present, such as oxygen, hydrogen, chloride, dichlorethene (DCE), methane, and others. The key model output is the probability that anaerobic degradation of trichloroethene (TCE) via reductive dechlorination is occurring. An expert judgment-based approach was used to estimate the output probability. A total of 14 possible pieces of "evidence" were considered in the analysis. The results of the *DLOR* analysis of Stiber *et al*. (1999) are given in Figure 1. For example, if there is no evidence that oxygen is present then there is increased probability of anaerobic processes occurring. If chloride or DCE is present, then there is positive evidence of the degradation of TCE. From the sensitivity analysis, it appears that evidence regarding DCE is the most important determinant of anaerobic degradation of TCE, because changes in evidence for this particular input lead to the largest differences in the log-odds ratio.

2.2.3 Advantages.

The ΔLOR method is a useful measure of sensitivity when the model output is a probability (Menard, 1995).

2.2.4 Disadvantages.

The ΔLOR method can only be used when the output is in terms of probability (Menard, 1995). It suffers from other drawbacks similar to nominal range sensitivity analysis. For example, similar to nominal range sensitivity analysis, $DLOR$ cannot account for nonlinear interactions between or among inputs. Similar to nominal range sensitivity analysis, the significance of differences among the sensitivities can be difficult to determine for nonlinear models and correlated inputs, making it potentially difficult to rank order key inputs.

2.3 Break-Even Analysis.

Break-even analysis is more of a concept than a specific method. Broadly speaking, the purpose of break-even analysis is to evaluate the robustness of a decision to changes in inputs (von Winterfeldt and Edwards, 1986).

2.3.1 Description.

Break-even analysis involves finding values of inputs that provide a model output for which a decision maker would be indifferent among the two or more risk management options. The combinations of values of inputs for which a decision maker is indifferent to the decision options are known as switchover or break-even values. Then, in order to assess the robustness of a choice between the options, one can evaluate whether the possible range of values of the model inputs corresponds with only one of the two choices (Morgan and Henrion, 1990). Indifference of a decision maker to the two choices is often represented by a break-even line or indifference curve such as an iso-risk curve. Ambiguity regarding selecting a particular choice exists if the uncertainty range associated with an output may correspond to either of the two or more possible choices. Different options that result in equivalent levels of risk reduction also can be identified so that a decision maker can evaluate these options. If there are more than two decision options, the analysis can get complex (von Winterfeldt and Edwards 1986).

2.3.2 Application.

Break-even analysis is often used in economics for purposes such as budget planning (Dillon, 1993). Break-even analysis has also found applications in several other fields such as health care (Boles and Fleming, 1996). Kottas and Lau (1978) describe the concept of stochastic break-even analysis. Starr and Tapiero (1975) explain the use of break-even analysis with consideration of risk.

As an example of a breakeven analysis, consider a choice between two medical treatment options: medication versus an operation. The patient's valuation of the possible outcome of each option is represented with a utility function, and the probability of success of the operation is not known. Figure 2 illustrates the combination of values for the utility of the medication and the probability of success of the operation that lead to a preference for medication, as in the upper left portion of the graph, or for the operation, as in the lower right portion of the graph. Along the indifference line, there is equal preference for both options. If the patient has a high utility for the outcome of the medication, and believes that the probability of success of the operation is low, then medication is clearly preferred. If the range of uncertainty in the probability of success is large, and if the range of uncertainty encloses the indifference line, then it will be ambiguous as to what treatment option is preferred. In the former case, a decision can be made that is robust to uncertainty. In the latter case, more information is needed to reduce uncertainty in one or more of the inputs in order to make a decision with a high degree of confidence.

2.3.3 Advantages.

The switchover or break-even point guides further modeling and elicitation. If the range of uncertainty regarding an input encloses the break-even point, then that input will be important in making a decision; that is, there will be uncertainty regarding which decision to take. In such a situation, further research can be directed so as to help the decision maker to narrow the range of uncertainty and make a decision with more confidence. On the other hand, if the uncertainty regarding an input does not enclose the break-even point then there will be high confidence regarding the decision.

2.3.4 Disadvantages.

Break-even analysis is not a straightforward method to apply. Though it is a useful concept, its application is increasingly complex as the number of sensitive inputs increases (von Winterfeldt and Edwards, 1986). There also is not a clear ranking method to distinguish the relative importance of the sensitive inputs.

2.4 Automatic Differentiation Technique.

The automatic aifferentiation (AD) technique is an automated procedure for calculating local sensitivities for large models. In AD, a computer code automatically evaluates first-order partial derivatives of outputs with respect to small changes in the inputs. The values of partial derivatives are a measure of local sensitivity.

2.4.1 Description.

Most existing sensitivity analysis methods based upon differentiation, such as numerical differential methods, have one or more of the following limitations: inaccuracy in the results, high cost in human effort and time, and difficulty in mathematical formulation and computer program implementation (Hwang *et al.*, 1997). To overcome these limitations, AD techniques were developed. AD is a technique to perform local sensitivity analysis and not a new method in itself. In AD the local sensitivity is calculated at one or more points in the parameter space of the model. At each point, the partial derivatives of the model output with respect to a selected number of inputs is evaluated.

AD is implemented by pre-compilers that analyze the code of the complex model and then add instructions needed to compute first or higher order derivatives in an efficient manner to save computational time and reduce complexity (Kedem, 1980; Rall, 1980; Carmichael *et al.*, 1997). The resulting expanded code is then complied with a standard compiler so that code can evaluate function values used in the model, outputs, and derivatives. Automatic Differentiation in Fortran (ADIFOR) is one of the softwares used for implementation of AD (Bischof *et al.*, 1992, 1994 and 1996).

2.4.2 Application.

Automatic differentiation finds application in models that involve complex numerical differentiation calculations such as partial derivatives, integral equations, and mathematical series (Hwang *et al.*, 1997). It is used in fields such as air quality (Carmichael *et al.*, 1997), aerodynamics (Issac and Kapania, 1997), mechanical structures (Ozaki *et al.*, 1995), and others. Christianson (1999) used AD for verification of part of a model.

As an example, Carmichael *et al*. (1997) applied ADIFOR to calculate the sensitivity of ambient air ozone concentration to air quality model inputs representing initial conditions, reaction rate constants, and others. For this purpose, dimensionless sensitivity coefficients were calculated based upon the numerical partial derivative of ozone concentration with respect to a selected input, divided by the ratio of the values of the concentration and the input:

$$S_{i,j} = \frac{\left(\dfrac{\Delta C_i}{\Delta \alpha_j}\right)}{\left(\dfrac{C_i}{\alpha_j}\right)} \qquad (2\text{-}1)$$

where,

$S_{i,j}$ = Normalized local sensitivity coefficients for i^{th} chemical specie and j^{th} input;

ΔC_i = The absolute change in the output concentration of i^{th} chemical specie;

$\Delta \alpha_j$ = The absolute change in the j^{th} input;

α_j = Values of model inputs; and

C_i = Concentration of a chemical specie i at a given time.

The normalized local sensitivity coefficients were calculated for a 5-day simulation for two different meteorological scenarios. The estimated ozone concentration was found to be most sensitive to the initial ozone concentrations in both scenarios, with normalized sensitivity values of 0.082 and 0.054 for the marine and land scenarios, respectively. In contrast, the results were relatively insensitive to the initial NO concentration in both cases, as revealed by a normalized sensitivity coefficient of approximately 0.001. Lumped sensitivities were also reported to give the global effect of a perturbation in a given input on the whole system. Lumped sensitivity of a chemical specie is the summation of normalized local sensitivities at different time stages for that chemical specie.

2.4.3 Advantages.

AD techniques, such as ADIFOR can be applied without having detailed knowledge of the algorithm implemented in the model. ADIFOR does everything automatically once it is appended with the main code (Bischof *et al.*, 1992). 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). Hwang *et al.* (1997) observed CPU time saving of 57% by using AD for sensitivity analysis as compared to using a traditional method. If the model is constantly modified and improved, then ADIFOR provides a convenient tool to easily accommodate such necessary model changes, which can be very difficult to do in the case of other techniques (Carmichael *et al.*, 1997).

2.4.4 Disadvantages.

The availability of AD technique may be limited to specific computer languages such as FORTRAN in the case of ADIFOR, requiring the user to provide FORTRAN code for the model. Because AD is a local technique, it suffers from the limitations of nominal range sensitivity analysis. Furthermore, unlike nominal range sensitivity analysis, the possible range of values of the inputs is not considered. The accuracy for sensitivity results is conditioned upon the numerical method used in the AD software. Also, for nonlinear models, the significance of differences in sensitivity between inputs is difficult to determine, making the rank ordering of key inputs potentially difficult. This method cannot be used if partial derivatives cannot be evaluated locally.

2.5 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: (1) description of the relation between variables; (2) control of predictor variables for a given value of a response variable; and (3) prediction of a response based on predictor variables (Neter *et al.*, 1996; Sen and Srivastava, 1990).

2.5.1 Description.

A relation between inputs and the output should be identified prior to regression analysis based on techniques such as scatter plots or upon understanding of the functional form of the model. Methods, such as stepwise regression, can be used to automatically exclude statistically insignificant inputs. 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).

Regression analysis is most properly performed on an independent random sample of data. The effect of inputs on the output can be studied 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:

$$Y_i = \boldsymbol{b}_o + \boldsymbol{b}_1 X_{1,i} + \boldsymbol{b}_2 X_{2,i} + \ldots + \boldsymbol{b}_m X_{m,i} + \boldsymbol{e}_i \qquad (2\text{-}2)$$

where,

Y_i = i^{th} output data point for i^{th} input data points;

$X_{j,i}$ = i^{th} input data point for the j^{th} input;

$\boldsymbol{b}_j$ = regression coefficient for the j^{th} input; and

$\boldsymbol{e}_i$ = 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. For a linear model, the regression coefficient $\boldsymbol{b}_j$, 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 deviation of the prediction of the regression model from the actual can be measured using the coefficient of multiple determination, 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 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). Neter *et al*. (1996) has suggested a transformation called correlation transformation which standardize the data such that the standardized regression coefficients fall between -1 and 1. The advantage of this transformation is that the round-off errors can be minimized and all regression coefficients have the same unit and hence, regression coefficients can be compared on equal basis

In the case of forward stepwise regression analysis, the incremental change in the R^2 values is an indication of the significance of sensitivity of the output to each newly introduced input. However, in the case of dependent or correlated inputs, the problem of multicollinearity can affect the robustness of the results of regression analysis (Neter *et al.*, 1996). A few types of nonlinear models can be transformed into linear models in the case of logistic regression (Hosmer and Lemeshow, 1989). Generalized linear models (GLM) provide flexibility to use correlated input data and non-normal error distributions. Logistic regression and Poisson regression are examples of GLM (McCullagh and Nelder, 1989; Searle, 1987).

2.5.2 Applications.

Regression analysis as a sensitivity analysis method is applied in various fields such as veterinary science (Taylor *et al.*, 1992), social sciences (Barrett *et al.*, 1986), food sciences (Cliff *et al.*, 1995), and food safety (Tienungoon *et al.*, 2000). McCarthy *et al.* (1995) used logistic regression for sensitivity analysis of a stochastic population model using logistic regression.

As one example, Helton *et al.* (1995) used stepwise regression analysis for sensitivity analysis of a model for contamination of food pathways associated with a severe accident at a nuclear power station. A probabilistic analysis was done and forward stepwise regression analysis was used to help identify which of the input distributions contributed most to uncertainty in the mean population dose. The importance of inputs was indicated by the order in which inputs entered the regression model, the changes in R^2 values with the entry of successive inputs in the regression model, and the standardized regression coefficients in the final regression model. The stepwise regression analysis for an example output, the mean population dose, is shown in Table I. The mean population dose is dominated by the variables PSMCS134, TFMCS, and PSMI131. Collectively these three variables account for 86 percent of the observed variation. The remaining seven variables contribute only an additional eight percent increase in the explanatory power of the model, and thus are relatively insensitive. Regression coefficients for all these variables were statistically significant and variables with insignificant regression coefficients were excluded from the forward stepwise regression.

2.5.3 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).

2.5.4 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 sample regression analysis can be critically dependent upon the selection of a functional form for the regression model (Neter *et al.*, 1996). 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, regression 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.

2.6 Analysis of Variance.

Analysis of variance (ANOVA) is a model independent 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 no assumption is needed regarding the functional form of relationships between inputs and the outputs. Furthermore, categorical inputs and groups of inputs can be addressed.

2.6.1 Description.

Inputs are referred to as "factors" and values of factors are referred to as factor levels in ANOVA. An output is referred to as a "response variable." 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. Quantitative factor levels are continuous quantities, such as temperature or amount of pathogens. Neter *et al.* (1996) and Winter *et al.* (1991) give detail applications and terminology for ANOVA.

ANOVA is a nonparametric method that is used to determine if values of the output vary in a statistically significant manner associated with variation in values for one or more inputs. If the output does not have a significant association with variation in the inputs, then the variation in the output is random. The exact nature of the relationship between the inputs and the output is not determined by ANOVA. Although the F-test is generally used to evaluated the significance of the response of the output to variation in the inputs, additional tests such as the Tukey test and Scheffé test can also be used, such as to evaluate the effect of different input value ranges (Montgomery, 1997; Hochberb and Tamhane, 1987).

In ANOVA, it is assumed that the output is normally distributed. 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).

2.6.2 Application.

ANOVA finds broad application across various fields such as health risk assessment (*e.g.*, Pet-Armacost *et al.*, 1999; Gerken *et al.*, 2000), material testing (*e.g.*, Golinkin *et al.*, 1997), food quality (Carlucci *et al.*, 1999, Ashraf *et al.*, 1999), microbiology (*e.g*., McElroy *et al.*, 2000), and microbial risk assessment (*e.g*., Marks *et al.*, 1998).

As an example of multifactor ANOVA, Carlucci *et al.* (1999) applied ANOVA to evaluate the effects of factors, such as the age at slaughter, storage temperature, and the storage time on response variables such as odor, flavor, and texture of the lamb meat. Each response variable was evaluated using a 0–100 scale for different combinations of the inputs. The importance of input values was evaluated in terms of the relative magnitude of F values. All of the response variables were significantly affected by the age at slaughter. In addition, the response variable for texture of the meat was also influenced by an interaction term, indicating that the effect of storage time depends on the age at slaughter.

2.6.3 Advantages.

No assumption is needed regarding the type of underlying model and both continuous and discrete inputs can be analyzed using ANOVA (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 multicollinearity.

2.6.4 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 nominal range sensitivity analysis, 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.

2.7 Response Surface Method (RSM).

The Response Surface Method (RSM) can be used to represent the relation between a response variable (output) and one or more explanatory inputs (Myers and Montgomery, 1995; Neter *et al*., 1996; Khuri and Cornell, 1987). The RSM can be used in a probabilistic analysis (Chun *et al.*, 1996; Frey and Bharvirkar, 1998). The RSM can identify curvatures in the response surface by accounting for higher order effects. The RSM is generally complex and therefore, used in later stages of an investigation when a limited number of factors are under investigation (Neter *et al.*, 1996).

2.7.1 Description.

A Response Surface (RS) can be linear or nonlinear and is typically classified as first-order or second-order (Myers and Montgomery, 1995). The second-order structure is used when there are interactions terms between inputs. The amount of time and effort required to develop a response surface is typically a function of the number of inputs included and the type of response surface structure required. It is often advantageous to limit the number of inputs that are included in the response surface to those that are identified as most important using a screening sensitivity analysis method, such as nominal range sensitivity analysis. The model must be exercised for various desired combinations of the selected input values in order to generate a data set that can be used to fit or calibrate a response surface. Monte Carlo simulation methods are typically used to generate multiple values of each model input and to calculate corresponding values of the model output.

A typical approach to response surface development is to use a least-squares regression method to fit a standardized first or second order equation to the data obtained from the original model. The key assumptions of least-squares regression (*e.g.*, normality of the residuals) should be reasonably satisfied; otherwise, other techniques such as rank-based or nonparametric approaches, should be considered (Khuri and Cornell, 1987; Vidmar and McKean, 1996). The precision and accuracy of the response surface can be evaluated by comparing the predictions of the response surface to that of the original model for the same values of the model inputs. If the precision and accuracy is not satisfactory, an improved fit might be obtained by iterating on values of parameters for the response surface (Gardiner and Gettinby, 1998).

Once a response surface has been developed, the sensitivity of the model output to one or more of the selected inputs can be determined by: (1) inspection of the functional form of the response surface; (2) statistical analysis if regression analysis was used to develop the response surface; or (3) application of other sensitivity analysis methods to the response surface. The response surface can be thought of as a "model of a model" with the advantage of being simpler and faster to execute than the original model. Therefore, computationally intensive sensitivity analysis methods, such as Mutual Information Index or others, may be more readily applicable to the response surface than to the original model.

2.7.2 Application.

The RSM is often employed for optimization and product quality studies (Gardiner and Gettinby, 1998). Moskowitz (1997) demonstrated the use of RSM to optimize properties of cereals to maximize consumer acceptance. Hopperstad *et al.* (1999) applied RSM for reliability analysis of an aluminum extrusion process. The RSM also can be used for sensitivity analysis of stochastic systems (Williams *et al.*, 1999).

For example, Williams *et al.* (1999) describe application of RSM for optimization and sensitivity analysis of manufacturing process of a type of integrated circuit. The complex computations required for simulation of the manufacturing process previously limited the strategies for process optimization. Therefore, a RS was incorporated into the simulator software to enable statistical optimization and to perform sensitivity analysis. The RS included five of the twenty original model inputs and was second order in structure.

Frey and Bharvirkar (1998) used the RSM to simplify a simulation model of a coal gasification system. The RS models were developed from the data generated by probabilistic simulation of the gasification process. Using regression techniques, a RS model was fitted to the data. A series of equations were developed to enable prediction of 60 model outputs based upon combinations of 12 inputs. These RS models were implemented with a FORTRAN-based capital, operating, and levelized cost model to enable simultaneous prediction of the performance, emissions, and cost of the complex technology of gasification using desktop computers.

2.7.3 Advantages.

A key advantage of the RSM approach is that a potentially computationally-intensive model can be reduced to a simplified form that enables much faster model run times. Therefore, it will be easier to apply iterative numerical procedures to the RS, such as optimization or Monte Carlo simulation, compared to the original model. Furthermore, the functional form of the RS model and the values of its coefficients may provide a useful indication of key sensitivities. Nominal range sensitivity or other methods can be applied to the RS model to elucidate sensitivities, with faster run times.

2.7.4 Disadvantages.

In order to develop a RS, calculations with the original model are needed, which can be resource intensive for some models. Because the RS is calibrated to data generated from the original model, the valid domain of applicability of the RS model will be limited to the range of values used to generate the calibration data set. Most RS studies are based on fewer inputs than the original model. Therefore, the effect of all original inputs on the sensitivities cannot be evaluated in RSM.

2.8 Fourier Amplitude Sensitivity Test.

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.

2.8.1 Description.

The main feature of the FAST method is a pattern search method that selects points in the input parameter space, and which is reported 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 (Cuckier *et al.*, 1973). The contribution of input x_i to the total variance is calculated based on the Fourier coefficients, fundamental frequency w_i, and higher harmonics of the frequency as explained by Cuckier *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 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 adversely affected. 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).

2.8.2 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 inputs produced by the FAST method. The ranking of top three inputs was found to be robust but the FAST method could not consistently rank other inputs of the set.

2.8.3 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.

2.8.4 Disadvantages.

The FAST method suffers from computational complexity for a large number of inputs (Saltelli and Bolado, 1998). The classical FAST method is applicable to 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).

2.9 Mutual Information Index.

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

2.9.1 Description.

The MII method typically involves three general steps: (1) generating an overall confidence measure of the output value; (2) obtaining a conditional confidence measure for a given value of an input; and (3) calculating sensitivity indices (Critchfield and Willard, 1986). The overall confidence in the output is estimated from the CDF of the output. Confidence is the probability for the outcome of interest. For example, if the dichotomous output is whether risk is acceptable, the confidence is the probability that the risk is less than or equal to an acceptable level. Conditional confidence is estimated by holding an input constant at some value and varying all other inputs. The resulting CDF of the output indicates the confidence in the output conditioned on a particular value of the input. The mutual information between two random variables is the amount of information about a variable that is provided by the other variable

(Jelinek, 1970). The MII for each input is calculated based on the distribution of the input and on both the overall and conditional confidence in the output. The average mutual information index for an input is given by (Critchfield and Willard, 1986):

$$I^a_{XY} = \Sigma_x \Sigma_y P_X P_{Y|X} \log_n (P_{Y\,|\,X} / P_Y) \qquad (2\text{-}3)$$

where,

$P_{Y|X}$ = conditional confidence;

P_Y = overall confidence;

P_X = probability distribution for the input; and

n = 2, to indicate binary output.

I^a_{XY} is always positive. If I^a_{XY} is large, then X provides a great deal of information about Y. If X and Y are statistically independent, then I_{XY} is zero. The amount of information about a variable that is provided by the variable itself is measured in terms of the "average self information" (I_{YY}) of that variable, also known as the entropy of that variable (Jelinek, 1970). The Average Self Information of the output Y is given as:

$$I_{YY} = \Sigma_Y P_Y \log_n (1 / P_Y) \qquad (2\text{-}4)$$

where P_Y is the overall confidence measure of Y. For the purpose of sensitivity analysis, a normalized measure of the MII, the "natural sensitivity index," S_{XY} is used.

$$S_{XY} = (I^a_{XY} / I_{YY}) \text{ X } 100\ \% \qquad (2\text{-}5)$$

S_{XY} reflects the percentage of the average mutual information contribution to the model output Y that can be attributed to the input variable X. The calculation of MII may require simplifying approximations regarding the use of a limited number of input values to represent the variation in an input and estimation of probabilities of the input values. No analytical statistical measure is available to determine the significance of the sensitivity indices. Because of the assumptions and simplifications made in evaluating S_{XY}, the robustness of rank ordering of key inputs can be difficult to evaluate.

2.9.2 Application.

The MII method was devised by Critchfield and Willards (1986), who demonstrated its application on a decision tree model. A dichotomous model was used to decide between two options to treat the disease of deep vein thrombosis (DVT): anticoagulation and observation. Each of these options had a relative importance to the decision maker, and they were valued in terms of utility. The difference in the utilities (ΔU) between the two options was used determine the preferred option. The probability of a particular value of ΔU for an option was interpreted as the overall confidence level for that option. The overall confidence level for anticoagulation was found to be 0.92 (92%).

In the conditional confidence analysis, an input of interest, such as the probability of pulmonary embolism (PE) given DVT, P(PE|DVT) was held constant at one value and other inputs were varied according to their respective distributions. The conditional confidence on ΔU for anticoagulation was determined. This process was repeated for all values of the input. The results of the conditional confidence analyses were represented graphically. The mutual information index and natural sensitivity index, S_{XY}, were calculated. S_{XY} for the input variable, P(PE|DVT) was 25. The input variable, P(fetopathy), had a sensitivity index of 3.6. These results indicated that the model output ΔU was more sensitive to P(PE|DVT) than P(fetopathy).

2.9.3 Advantages.

MII includes the joint effects of all the inputs when evaluating sensitivities of an input. The mutual information is a more direct measure of the probabilistic relatedness of two random variables than other measures such as correlation coefficients (Jelinek, 1970). For example, the correlation coefficient of two random variables examines the degree of linear relatedness of the variables. Although two uncorrelated variables may not be independent, two variables with zero mutual information are statistically independent. Therefore, the MII is a more informative method. The results can be presented graphically facilitating their comprehension.

2.9.4 Disadvantages.

Calculation of the MII by Monte Carlo techniques suffers from computational complexity, making practical application difficult (Merz *et al.*, 1992). Critchfield and Willard (1986a) have suggested an approach using symbolic algebra, which is reported to be less computationally intensive. Because of the simplifying approximations that may be used in MII, the robustness of ranking based on the sensitivity measure can be difficult to evaluate.

2.10 Scatter Plots.

A scatter plot is a graphical sensitivity analysis method. 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.

2.10.1 Description.

Each realization in a probabilistic simulation, such as a Monte Carlo simulation, generates one pair of an input value and the corresponding output value. These simulated pairs can be plotted as points on a scatter plot. Scatter plots depict the possible dependence between an input and the output. Dependence may be linear or nonlinear (Cook, 1994). The range of variation of the output may be constant regardless of the specific value of the input, or it may be non-constant. Scatter plots may be based upon sample values, or based upon the ranks of the values. General trends that might be observed in scatter plots are described by Kazmierski (1995) and Bernstein *et al*. (1988).

Because scatter plots can help in visualizing and identifying potentially complex dependencies between an input and an output, they can be used to guide the selection of appropriate sensitivity analysis methods. For example, if the relationship is nonlinear, then a nonlinear regression model, or a transformation of the data, may be required.

The number of data points displayed in a scatter plot needs to be selected such that there is enough density to be able to observe any pattern, but not so many points that the variability within the scatter is difficult to observe (Vose, 2000).

Kleijnen and Helton (1999) describe a sensitivity analysis method based on pattern detection used to identify relationships between inputs and an output. For example, pattern detection methods can identify linear and monotonic relationships between inputs and an output as well as the effect of marginal distributions of an input and an output. Shortencarier and Helton (1999) describe a pattern detection method and a related software tool.

2.10.2 Application.

Scatter plots have been used as an aid to sensitivity analysis in various fields such as behavioral studies (*e.g.*, Sobell, 1982; Rossier *et al.*, 2001), environmental pollution (*e.g.*, Hagent, 1992; Fujimoto, 1998), plant safety (*e.g.*,Helton *et al.*, 2000), medical science (*e.g.*, Bruno et al., 1994; Kogo and Ariel, 1997), veterinary science (*e.g.*, Glaser, 1996), and others.

As an example, Moskowitz (1997) used consumer-based product evaluation to identify optimal product formulations for a ready-to-eat cereal. Several physical attributes of the cereal, such as appearance, color, flavor and quality which depend upon parameters (or inputs) such as the amount of sweeteners and starch, die size, roasting time, and others, impact the customer liking of the cereal. Thirty-one different product mixes were prepared by using different combinations of eight such parameters. Different panels of consumers and experts tested these mixes and used a 0 to 100 scale to rate their liking for the attributes of the products. Scatter plots were used to represent the relationship between different combinations of the parameters (or inputs) and the overall ratings of the cereal attributes (output).

2.10.3 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. An understanding of the nature of the dependencies between inputs and an output can guide the selection of other appropriate sensitivity analysis methods.

2.10.4 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 (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.

3.0 COMPARISON OF METHODS.

In this section, each of the ten methods are compared based upon four selected criteria: (1) applicability of the method; (2) computational intensiveness; (3) ease and clarity in representation of sensitivity; and (4) the purpose of the analysis. A summary of the comparison for all ten methods is provided in Table II.

3.1 Applicability for Different Types of Models.

As indicated in Table II, some methods can be used only with deterministic models, while others require the availability of probabilistic simulation results. Furthermore, some methods are predicated upon an assumption of linearity, such as NRSA. Similarly, although AD can be applied at multiple points in the input variable space, at each specific point it provides a linear measure of sensitivity. Some methods are model dependent, such as regression analysis and response surface methods, while the other methods do not require any assumption regarding the functional form of the model. The ΔLOR method and the AD method have specific requirements regarding the form of the model output and regarding differentiability of the model, respectively. The FAST method can be challenged by the presence of discrete inputs.

FSRMs are nonlinear, contain thresholds, and typically contain discrete inputs. At the same time, FSRMs typically include probabilistic simulations of variability and/or uncertainty. These characteristics suggest potential problems in application of NRSA and AD, which are typically deterministic methods that assume linearity, and FAST, which may have difficulty with discrete inputs in some cases. The large number of inputs to a FSRM may make break-even analysis impractical; however, break-even could be applied to a FSRM after a small number of most sensitive inputs have been identified using other methods. Methods such as ANOVA and MII appear to be the most theoretically attractive methods, because they are model independent and take into account the simultaneous interaction of all selected model inputs. However, MII in particular is computationally intensive. Thus, it may be useful to develop a simplified response surface based upon key inputs identified using less computationally intensive methods, and then apply MII to the response surface.

The accuracy of a sensitivity analysis method will depend on whether all key assumptions of the method hold true for a particular application. For example, if NRSA is used with a model that is not only nonlinear, but also non-monotonic, it may be possible to obtain the wrong insight. As an illustration, suppose that the value of a model output is the same for both the low and high values of a given model input, but that the output has a maximum in the middle of the range of values for the model input. NRSA will not reveal any relationship in this case. Moreover, in the absence of any other analyses or insights by the analyst, NRSA will not by itself reveal its own shortcoming in this case. In contrast, there are standard diagnostic tests that can be used with methods such as regression analysis and ANOVA to determine if key assumptions have been violated.

3.2 Computational Issues.

FSRMs can be large and computationally intensive. Therefore, there are incentives to identify sensitivity analysis methods that do not introduce substantial additional computational burdens. Some of the methods identified, including regression analysis, ANOVA, and scatter plots, can be applied to the results of a probabilistic analysis without need for any additional model runs. This can be useful only if all candidate sensitive variables have been assigned distributions in the probabilistic analysis. Moreover, computational intensity is based not just on exercise of the original model, but also upon any additional computations required in applying the sensitivity analysis method. For example, the time to perform an ANOVA calculation will be sensitive to the sample size, the number of inputs, and whether there are interactions among the inputs. The time required to prepare scatter plots will depend on the number of pair-wise combinations of inputs and outputs that are evaluated. Commercial software are readily available in both of these examples, as well as for other methods such as regression analysis.

NRSA is potentially an easy method to apply if a limited number of inputs are to be evaluated. In a FSRM, there may be some inputs that are not treated probabilistically which could potentially be important in determining the model output. NRSA is in some ways preferred over AD methods, because both the sensitivity and possible range of values are taken into account in the former, while only local sensitivity is considered in the latter.

Aside from computational intensity, another consideration is computational feasibility. AD methods can only be applied if the model is locally differentiable. FAST can encounter difficulties with discrete model inputs. Thus, it is possible that either of these two methods may be inapplicable to a particular model, unless the model or method can be revised to correct for these issues.

3.3 Ease and Clarity in Representation of Sensitivity.

Each sensitivity analysis method has its own metric for sensitivity. There are standard numerical ways of representing results for NRSA, ΔLOR, AD, regression analysis, ANOVA, FAST, and MII. For regression and ANOVA based methods, there are statistical measures of the significance of the sensitivity indicators (*e.g.*, regression model coefficients, F-ratios) that can be used to help distinguish which inputs are insensitive from those that have at least some influence on the output. Regression-based methods also provide information that might help distinguish among the statistically significant inputs. For example, in standardized step-wise regression, the magnitude of the regression coefficient and the incremental improvement in the R^2 value can be compared for different inputs to gain insight into which ones are substantially more sensitive than others. In some cases, a model output may have approximately equal sensitivity to two or more inputs. Thus, small differences in the sensitivity indicators may be insignificant from a practical, if not statistical, perspective. As a result, there may be ambiguity regarding the rank order of inputs from most sensitivite to least sensitive with respect to a given output.

For the other methods, such as break-even analysis, RSM, and scatter plots, the representation of sensitivity may be less prescribed and more subject to individualized presentation and interpretation. In the case of RSM, the interpretation of sensitivity may be based upon examination of the functional form of the response surface, and the values of its coefficients. If a statistical method, such as regression, was used to develop the response surface, then there may be statistical measures of the significance of sensitivities between the output and one or more inputs (including interaction terms). The interpretation of break-even analysis results typically distills to whether or not a decision is sensitive to the range of uncertainty in one or more inputs. Additional analysis might clarify how sensitive the decision may be. For example, one could estimate the probability that a particular option is not the best one depending on the distribution of uncertainty assigned to one of the key decision inputs. In the case of scatter plots, the interpretation will often be case-specific. In some situations, clear insights may emerge that enable identification of the most important inputs.

Sensitivity analysis methods can be distinguished based upon whether they are local or global methods, and regarding whether they focus on one input at a time or take into account simultaneous interactions among the inputs. A method that is global and that takes into account interactions among multiple inputs may provide more robust insights than methods without these features. NRSA, ΔLOR, and AD are local methods and they do not account for simultaneous interactions among the inputs. Break-even analysis may account for interactions among a small number of inputs, conditioned on either point-estimates or distributions for all remaining inputs. Thus, break-even analysis is a hybrid of both local and global characteristics. Regression, ANOVA, FAST, MII, RSM, and scatter plots can be applied over a potentially large domain of the input parameter space and can be applied taking into account simultaneous variation of multiple inputs.

3.4 Purpose of the Analysis.

Sensitivity analysis may be performed for a variety of reasons, including identification of key inputs and/or exploration of the model response to specific inputs. The latter can be useful to either verification and/or validation of the model.

All of the methods reviewed here have the potential to reveal key sensitivities as long as key assumptions for a given method are not grossly violated. Some methods provide more information regarding the nature of the sensitivity than others. For example, because ANOVA is a model-free and nonparametric approach, it only provides indication of a statistical association between an output and an input. MII provides a measure of the information that an input imparts to the output. FAST provides insight regarding the contribution of an input to the variance in the output. None of these methods provide insight into the specific nature of the relationship. Thus, these methods may be helpful in identifying key inputs, and in identifying qualitative features of the relationship between an input and an output, but they are less useful in terms of quantitative verification and/or validation of a model.

In contrast, a technique such as regression analysis or RSM can be used to help find a key functional relationship that enables quantitative estimation of how the output changes with respect to variation in the input. Methods such as NRSA, ΔLOR, and AD provide insight regarding the local behavior of the model, and thus have the potential to be useful for both verification and validation. Scatter plots may in some sense be the most revealing of possible problems in model performance, since no prior assumption is made regarding how the model is responding to a given input.

Of course, with simple modifications, some methods can have increased utility as a means for verification or validation of model performance. For example, a strict application of NRSA may not take into account nonlinearities in model response. However, if three or more values of a given model input are used to predict the model output, and if all such values are displayed graphically, any nonlinearity in the model response can be revealed and compared with prior expectations or with validation data.

4.0 CONCLUSIONS.

There are a large number of sensitivity analysis methods used in a variety of disciplines. Only ten such methods have been identified and evaluated here. The methods were selected based upon a judgment that they are widely used and of potential relevance to FSRMs. Each method was characterized individually, and the methods were compared on the basis of four criteria.

There is no single method that is clearly superior to all others. Each method has its own key assumptions and limitations. Each method has its own demands regarding the time and effort to apply the method and interpret the results. Each method has strengths and limitations regarding the type of insight that it can provide. FSRMs have important features that, taken individually, favor one method over another. However, when taken in aggregate, there is not one obvious best method. For example, a model-free and global method is recommended as the most appropriate choice because FSRMs typically are nonlinear, contain thresholds, and include discrete inputs. Thus, methods such as NRSA, ΔLOR, AD, regression analysis, and RSM may not be able to provide robust insights either because they assume linearity or because they require specification of a functional form. FAST, in spite of its many attractive features, may be inapplicable because of potential problems in dealing with discrete inputs. In contrast, methods such as ANOVA, MII, and scatter plots are model independent and in principle can deal with the key characteristics of FSRMs. ANOVA can be applied to the results of a probabilistic analysis but does not by itself provide direct insight into the nature of the relationship between the output and the most sensitive inputs. MII would require computationally intensive simulations which may be impractical unless a good response surface can be used instead of the original model. Scatter plots should be used to help identify relationships that may otherwise be difficult to characterize with other methods.

An iterative approach to sensitivity analysis is likely to be needed for most FSRMs. In the iterative approach, one or more methods that are not computationally intensive can be used to make a preliminary identification not only of the most important sensitive inputs, but also of the least important inputs. Subsequently, more refined and/or computationally intensive sensitivity analysis methods can be applied only to the subset of inputs that appear to be most important or that a particular method can easily address.

Because each sensitivity analysis methods is typically based upon a different assumption regarding appropriate ways of measuring sensitivity, it is possible that two different methods may lead to different rank ordering of key inputs. Thus, a general recommendation is to use two or more methods, preferably with dissimilar foundations, to increase confidence that the identification of key inputs is robust.

Other important issues that must be addressed with FSRMs include modularity and aggregation. FSRMs typically consistent of separate modules that represent different food processing steps. Within a given part of a model, there may be aggregation (*e.g.*, from individual animals to batches of ground beef). As part of future work, methods should be identified, applied, and evaluated for sensitivity analysis across modules and taking into account aggregation.

Although there are theoretical arguments in favor of some methods over others, these methods should be compared to evaluate whether their results differ in practice. Thus, for future work, a quantitative comparison of multiple sensitivity analysis methods applied to the same specific refined FSRMs is recommended. Such a comparison can provide insight regarding whether the methods, in spite of different theoretical foundations, perform similarly in practice.

5.0 ACKNOWLEDGEMENT.

This work was supported by a cooperative agreement between the U.S. Department of Agriculture's Office of Risk Asseessment and Cost Benefit Analysis and North Carolina State University. The authors benefited from suggestions provided by participants of the June 11-12, 2001 NCSU/USDA Workshop on Sensitivity Analysis and from peer reviewer comments.

6.0 REFERENCES.

Agro, K.E., C.A. Bradley, and N. Mittmann (1997), "Sensitivity Analysis in Health Economic and Pharmacoeconomic Studies - an Appraisal of the Literature," *Pharmacoeconomics*, 11(1):75-88.

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.

Ashraf, H.R., M. White, and B. Klubek (1999), "Microbiological Survey of Tofu Sold in a Rural Illinois County," *Journal of Food Protection*, 62(9):1050-1053.

Baker, S., D. Ponniah, and S. Smith (1999), " Survey of Risk Management in Major UK Companies," *Journal of Professional Issues in Engineering Education and Practice*, 125(3):94-102.

Baniotopoulos, C.C. (1991), "A Contribution to the Sensitivity Analysis of the Sea-Bed Structure Interaction Problem for Underwater Pipelines," *Computers & Structures*, 40(6):1421-1427.

Barrett, G.V., R.A. Alexander, M.N. Anesgart, and D. Doverspike (1986), "Frequently Encountered Problems in the Application of Regression-Analysis to the Investigation of Sex-Discrimination in Salaries," *Public Personnel Management*, 15(2):143-157.

Beck, M.B., J.R. Ravetz, L.A. Mulkey, and T.O. Barnwell (1997), "On the Problem of Model Validation for Predictive Exposure Assessments," *Stochastic Hydrology and Hydraulics*, 11(3):229-254.

Bernstein, I.H., C.P. Garbin, and G.K. Teng (1988). *Applied Multivariate Analysis*. Springer-Verlag: New York.

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, Aargonne National Laboratory, Argonne, Illinois.

Bischof, C.H., A. Carle, P. Khademi, and A. Mauer (1994), “The ADIFOR2.0 System for the Automatic Differentiation of FORTRAN77 Programs.” *IEEE Computational Science and Engineering*, 3(3):18-32.

Bischof, C.H., A. Carle, P. Khademi, and A. Mauer (1996), "ADIFOR 2.0: Automatic Differentiation of Fortran 77 Programs," *IEEE Computational Science & Engineering*, 3(1):18-32.

Boles, K.E., and S.T. Fleming (1996), "Break-even Under Capitation: Pure and Simple?," *Health Care Management Review*, 21(1):38-47.

Box, G.E.P., and G.C. Tiao (1992). *Bayesian Inference in Statistical Analysis*. John Wiley and Sons, Inc.: New York.

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.

Bruno, A., G. Delpoeta, A. Venditti, R. Stasi, G. Adorno, G. Aronica, G. Suppo, A.M. Dirienzo, R. Iazzoni, M. Tribalto, T. Caravita, M. Masi, and G. Papa (1994), "Diagnosis of Acute Myeloid-a System Coulter VCS," *Haematologica*, 79(5):420-428.

Buchanan, R.L. (1995), "The Role of Microbiological Criteria and Risk Assessment in HACCP," *Food Microbiology*, 12(5):421-424.

Buchanan, R.L. (1996), "Indentifying and Controlling Emerging Food Borne Pathogens: Research Need," *Emerging Infectious Diseases*, 3(4):517-521.

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

Carlucci, A., F Napolitano, A. Girolami, and E. Monteleone (1999), "Methodological approach to evaluate the effects of age at slaughter and storage temperature and time on sensory profile of lamb meat," *Meat Science*, 52(4):391-395.

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.

Cheng, T.C.E. (1991), "EPQ with Process Capability and Quality Assurance Considerations," *Journal of the Operational Research Society*, 42(8):713-720.

Christensen, Ronald (1990). *Log-Linear Models*. Springer-Verlag: New York.

Christianson, B. (1999), "Cheap Newton Steps for Optimal Control Problems: Automatic Differentiation and Pantoja's Algorithm," *Optimization Methods and Software*, 10(5):729-743.

Chun, M.H., S.J. Han, and N.I. Tak (1996), "A Combined Procedure of RSM And LHS for Uncertainty Analyses in Source Term Quantifications Using Maap3.0B Code," *Annals of Nuclear Energy*, 23(16):1337-1349.

Cliff, M.A., M.C. Dever, J.W. Hall, and B. Girard (1995), "Development and Evaluation of Multiple Regression Models for Prediction of Sweet Cherry Liking," *Food Research International*, 28(6):583-589.

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 Mote Carlo Simulation," *Medical Decision Making*, 6(1):85-92.

Critchfield, G.C., and K.E. Willard (1986a), "Probabilistic Analysis of Decision Trees Using Symbolic Algebra," *Medical Decision Making*, 6(1):93-100.

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

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.

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

Dakins, M.E., J.E. Toll, and M.J. Small (1994), "Risk-Based Environmental Remediation: Decision Framework and Role of Uncertainty," *Environmental Toxicology and Chemistry*, 13(12):1907-1915.

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

Dillon, C.R. (1993), "Advanced Break-Even Analysis of Agricultural Enterprise Budgets," *Agricultural Economics*, 9(2):127-143.

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

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

FAO (1999), "CODEX Alimentarius Commission," Report No. ALINORM 99/13A. Food and Agriculture Organization (FAO) of the United Nations. http://www.codexalimentarius.net/, 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.

Fontaine, T.A., and V.M.F. Jacomino (1997), "Sensitivity Analysis of Simulated Contaminated Sediment Transport," *Journal of the American Water Resources Association*, 33(2):313-326.

Fraedrich, D., A. Goldberg (2000), "A Methodological Framework for the Validation of Predictive Simulations," *European Journal of Operational Research*, 124(1):55-62.

Frey, H.C., and R. Bharvirkar (1998), "Desktop Modeling of the Performance, Emissions, and Cost of Gasification Systems," *Proceedings of the 91st Annual Meeting*, Paper No. 98-RAA.02P, Air and Waste Management Association, Pittsburgh, Pennsylvania.

Fujimoto, K. (1998), " Correlation Between Indoor Radon Concentration and Dose Rate in Air from Terrestrial Gamma Radiation in Japan," *Health Physics*, 75(3):291-296.

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.

Gardiner, W.P., and G. Gettinby (1998). *Experimental Design Techniques in Statistical Practice: A Practical Software-Based Approach*. Horwood Publishing Ltd.: Chichester, West Sussex, England.

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.

Gerken, M., L. Kreienbrock, J. Wellmann, M. Kreuzer, and H.E. Wichmann (2000), "Models for Retrospective Quantification of Indoor Radon Exposure in the Case-Control Studies," *Health Physics*, 78(3):268-278.

Glaser, S. (1996), "Application of Statistical Classification Methods to the Evaluation of In Vitro Tests," *Alternatives to Laboratory Animals*, 24(1):49-62.

Golinkin, I.A., D.D. Ruff, E.P. Kvam, G.P. McCabe, and A.F. Grandt (1997), "Application of Analysis of Variance (ANOVA) Statistical Methods to Breaking Load Corrosion Test," *Journal of Testing and Evaluation*, 25(6):565-570.

Greene, T., and C.B. Ernhart (1993), "Dentin Lead and Intelligence Prior to School Entry - A Statistical Sensitivity Analysis," *Journal of Clinical Epidemiology*, 46(4):323-339.

Hagent, N.T. (1992), " Macroparasitic Epizootic Disease: A potential Mechanism for the Termination of Sea-Urchin Outbreaks in Northern Norway," *Marine Biology*, 114(3):469-478.

Hathaway, C.N. (1995), "Harmonization Requirements Under HACCP-Based Control Systems," *Food Control*, 6(5):267-276.

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, J.C., J.D. Johnson, J.A. Rollstin, A.W. Shiver, and J.L. Sprung (1995), "Uncertainty and Sensitivity Analysis of Food Pathway Results With the MACCS Reactor Accident Consequence Model," *Reliability Engineering and System Safety*, 49(2), 109-144.

Helton, J.C., J.E. Bean, K. Economy, J.W. Garner, R.J. MacKinnon, J. Miller, J.D. Schreiber, P. Vaughn (2000), "Uncertainty and Sensitivity Analysis for Two-Phase Flow in the

Vicinity of the Repository in the 1996 Performance Assessment for the Waste Isolation Pilot Plant: Disturbed Conditions," *Reliability Engineering and System Safety*, 69(1-3):263-304.

Hochberg, Y., and A.C. Tamhane (1987). *Multiple Comparison Procedures*. John Wiley and Sons: New York.

Hoffman, F.O., and C.W. Miller (1983). "Uncertainties in Environment Radiological Assessment Models and Their Implications," Proceedings of the Nineteenth Annual Meeting of the National Council on Radiation Protection and Measurements. National Academy of Science, Washington, DC.

Hopperstad, O.S., B.J. Leira, S. Remseth, and E. Tromborg (1999), "Reliability-based Analysis of a Stretch-bending Process for Aluminum Extrusions," *Computers and Structures*, 71(1):63-75.

Hosmer, D.W., and S. Lemeshow (1989). *Applied Logistic Regression*. John Wiley: New York.

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.

Issac, J.C., and R.K. Kapania (1997), “Aero-elastic Sensitivity Analysis of Wings Using Automatic Differentiation,” *Journal of American Institute of Aeronautics and Astronautics*, 35(3):519-525.

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

Jones, R.N. (2000), "Analyzing the Risk of Climate Change Using an Irrigation Demand Model," *Climate Research*, 14(2):89-100.

Kaplan, S., and B.J. Garrick (1981), "On the Quantitative Definition of Risk," *Risk Analysis*, 1(1):11-27.

Kazmierski, T.J. (1995). *Statistical Problem Solving in Quality Engineering*. McGraw-Hill, Inc: New York.

Kedem, G. (1980), " Automatic Differentiation of Computer Programs," *ACM Transactions on Mathematical Software*, 6(2):150-165.

Kendall, M.G., and J.D. Gibbons (1990). *Rank Correlation Methods*. Fifth Edition. Charles Griffin: London.

Kewley, R.H., M.J. Embrechts, and C. Breneman (2000), "Data Strip Mining for the Virtual Design of Pharmaceuticals with Neural Networks," *IEEE Transactions on Neural Networks*, 11(3):668-679.

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

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.

Kleijnen, J.P.C. (1995), "Verification and Validation of Simulation-Models," *European Journal of Operational Research*, 82(1):145-162.

Kleijnen, J.P.C., and J.C. Helton (1999). "Statistical Analysis of Scatter Plots to Identify Important Factors in Large-Scale Simulations." SAND98-2202, Prepared by Sandia national Laboratories, Albuquerque, New Maxico.

Kleijnen, J.P.C., and R.G. Sargent (2000), " A Methodology for Fitting and Validating Metamodels in Simulation," *European Journal Of Operational Research*, 120(1):14-29.

Kogo, N., and M. Ariel (1997), "Membrane Properties and Monosynaptic Retinal Excitation of Neurons in the Turtle Accessory Optic System," *Journal of Neurophysiology*, 78(2):614-627.

Kottas, J.F., and H.S. Lau (1978), "Stochastic Break-Even Analysis," *Journal of the Operational Research Society*, 29(3):251-257.

Krishnaiah, P.R (1981). *Ananlysis of Variance*. Elsevier: New York.

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

Limat, S., M.C. Woronoff-Lemsi, E. Deconinck, E. Racadot, M. Jacquet, P. Herve, and J.Y. Cahn (2000), "Cost-Effectiveness of Cd34(+) Dose in Peripheral Blood Progenitor Cell Transplantation for Non-Hodgkin's Lymphoma Patients: a Single Center Study," *Bone Marrow Transplantation*, 25(9):997-1002.

Lindman, H.R (1974). *Analysis of Variance in Complex Experimental Designs*. W. H. Freeman & Co.: San Francisco.

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.

Manheim, L.M. (1998), "Health Services Research Clinical Trials: Issues in the Evaluation of Economic Costs and Benefits," *Controlled Clinical Trials*, 19(2):149-158.

Marks, H.M., M.E. Coleman, C.T.J. Lin, and T. Roberts (1998), "Topics in Microbial Risk Assessment: Dynamic Flow Tree Process," *Risk Analysis*, 18(3):309-328.

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

McCarthy, M.A., M.A. Burgman, and S. Ferson (1995), “Sensitivity Analysis for Models of Population Viability,” *Biological Conservation*, 73(1):93-100.

McCullagh, P., and J.A. Nelder (1989). *Generalized Linear Models*. Second Edition. Chapman and Hall: London.

McElroy, D.M., L.A. Jaykus, and P.M. Foegeding (2000), "Validation and Analysis of Modeled Predictions of Growth of Bacillus Cereus Spores in Boiled Rice," *Journal of Food Protection*, 63(2):268-272.

McKone, T.E. (1996), "Overview of the Risk Analysis Approach and Terminology: the Merging of Science, Judgment and Values," *Food Control*, 7(2):69-76.

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.

Menard, Scott (1995). *Applied Logistic Regression Analysis*. Sage Publications: Thousand Oaks, California.

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

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.

Montgomery, D.C. (1997). *Design and Analysis of Experiments*. Wiley and Sons Ltd.: New York.

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.

Moskowitz, H.R. (1997), “A Commercial Application of RSM for Ready-to-Eat Cereal,” *Food Quality and Preference*, 8(3):191-201.

Myers, R.H., and D.C. Montgomery (1995). *Response Surface Methodology: Process and Product Optimization Using Designed 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.

Notermans, S., G. Gallhoff, M.H. Zwietering, and G.C. Mead (1995), "The HACCP Concept: Specification of Criteria Using Quantitative Risk Assessment," *Food Microbiology*, 12(2):81-90.

NRC (1983). *Risk Assessment in the Federal Government: Managing the Process*. National Research Council. National Academy Press: Washington, DC.

Oh, B.H., and I.H. Yang (2000), "Sensitivity Analysis of Time-Dependent Behavior in PSC Box Girder Bridges," *Journal of Structural Engineering-ASCE*, 126(2):171-179.

Ozaki, I., F. Kimura, and M. Berz (1995), "Higher-Order Sensitivity Analysis of Finite-Element Method by Automatic Differentiation," *Computational Mechanics*, 16(4):223-234.

Pet-Armacost, J.J., J. Sepulveda, and M. Sakude (1999), "Monte Carlo Sensitivity Analysis of Unknown Parameters in Hazardous Waste Transportation Risk Assessment," *Risk Analysis*, 19(6):1173-1184.

Phillips, A., D. Janies, and W. Wheeler (2000), "Multiple Sequence Alignment in Phylogenetic Analysis," *Molecular Phylogenetics And Evolution*, 16(3):317-330.

Rall, L.B (1980). *Fundamentals of Numerical Computation: Applications of Software for Automatic Differentiation in Numerical Computation*. Computing Supplement No. 2. Springer Verlag: Berlin, New York.

Reed, J.A., A.A. Afjeh (2000), "Computational Simulation of Gas Turbines: Part 1 - Foundations of Component-Based Models," *Journal of Engineering for Gas Turbines and Power-Transactions of the ASME*, 122(3):366-376.

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.

Rossier, P., D.T. Wade, M. Murphy (2001), "An Initial Investigation of the Reliability of the Rivermead Extended ADL Index in Patients Presenting with Neurological Impairment," *Journal of Rehabilitation Medicine*, 33(2):61-70.

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., K. Chan, and E.M. Scott. Eds (2000). *Sensitivity Analysis*. John Wiley and Sons, Ltd.: West Sussex, England.

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.

Searle, S. R (1987). *Linear Models for Unbalanced Data*. John Wiley and Sons: New York.

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.

Shortencarier M.J., and J.C. Helton (1999). "A FORTRAN 77 Program and User's Guide for the Statistical Analyses of Scatter plots to Identify Important Factors in Large-Scale Simluations." SAND99-1058, Prepared by Sandia national Laboratories, Albuquerque, New Maxico.

Simon, H.A (1996). *The Sciences of the Artificial*. Third Edition. MIT Press: Cambridge, MA.

Sinicio, R., R.E. Muir, and D.S. Jayas (1997), "Sensitivity Analysis of a Mathematical Model to Simulate Aeration of Wheat Stored in Brazil," *Postharvest Biology and Technology*, 11(2):107-122.

Sobell, L.C., S. Maisto, A.M. Cooper, and M.B. Sobell (1982), "Comparison of Two Techniques to Obtain Retrospective Reports of Drinking Behavior from Alcohol Abusers," *Addictive Behaviors*, 7(1):33-38.

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

Song, F.J., A.M. Glenny, and D.G. Altman (2000), "Indirect Comparison in Evaluating Relative Efficacy Illustrated by Antimicrobial Prophylaxis in Colorectal Surgery," *Controlled Clinical Trials*, 21(5):488-497.

SPLUS (2000). *Guide to Statistics, Volume I and II*. Data Analysis Product Division, MathSoft Inc.: Seattle, WA.

Starr, M.K., and C.S. Tapiero (1975), "Linear Break-Even Analysis Under Risk," *Operational Research Quarterly*, 26(4):847-856.

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.

Taylor, B.C., J.L. Stott, M.A. Thurmond, and J.P Picanso (1992), "Alteration in Lymphocyte Subpopulations in Bovine Leukosis Virus-Infected Cattle," *Veterinary Immunology and Immunopathology*, 31(1-2):35-47.

Tienungoon, S., D.A. Ratkowsky, T.A. McMeekin, and T. Ross (2000), "Growth Limits of Listeria Monocytogenes As a Function of Temperature, pH, NaCl, and Lactic Acid," *Applied And Environmental Microbiology*, 66(11):4979-4997.

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

von Winterfeldt D, and Edwards W (1986) *Decision Analysis and Behavioral Research*. Cambridge, UK: Cambridge University Press, 1986: 399-405.

Vose, David (2000). *Risk Analysis: A Quantitative Guide*. 2nd Edition. John Willey & Sons, Ltd: New York.

Walpole, R.E., and R.H. Myers (1993). *Probability and Statistics for Engineers and Scientists*. 5th Edition. Macmillan Publishing Company: New York.

Ward, M.P., and T.E. Carpenter (1996), "Simulation Modeling of the Effect of Climatic Factors on Bluetongue Virus Infection in Australian Cattle Herds - 1. Model Formulation, Verification and Validation," *Preventive Veterinary Medicine*, 27(1-2):1-12.

Weiss, R (1996), "An Approach to Bayesian Sensitivity Analysis," *Journal of the Royal Statistical Society Series B-Methodological*, 58(4):739-750.

WHO (2001), "Microbial Risk Assessment in Food," World Health Organization. http://www.who.int/fsf/mbriskassess, Viewed: March, 21, 2001.

Williams, S., K. Varahramyan, and W. Maszara (1999), "Statistical Optimization and Manufacturing Sensitivity Analysis of 0.18 μm SOI MOSFETs," *Microelectronic Engineering*, 49(3-4):245-261.

Winter, B.J., D.R. Brown, and K.M. Michels (1991). *Statistical Principles in Experimental Design*. Third Edition. McGraw-Hill Inc.: New York.

Wotawa, G., A. Stohl, and H.KrompKolb (1997), "Estimating the Uncertainty of a Lagrangian Photochemical Air Quality Simulation Model Caused by Inexact Meteorological Input Data," *Reliability Engineering & System Safety*, 57(1):31-40.

Zwietering, M.H., and S.J.C. van Gerwen (2000), "Sensitivity Analysis in Quantitative Microbial Risk Assessment," *International Journal of Food Microbiology*, 58(3):213-221.

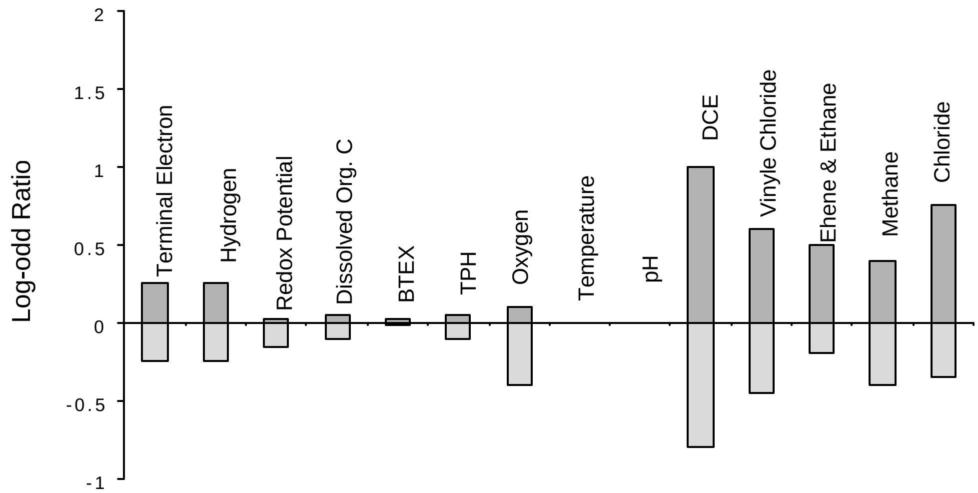

Figure 1. Example of chart for representing difference in log odds ratio.

(*Source:* Stiber *et al.*, 1999)

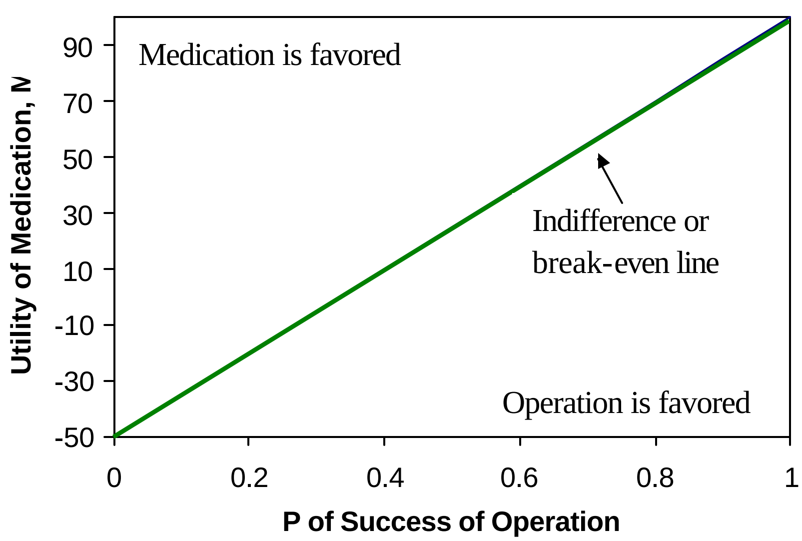

Figure 2. Example of break-even analysis.

Table I. Forward stepwise regression analysis for mean population dose.

(*Source:* Helton *et al.*, 1995)

Steps	**Inputs**	**Standardized Regression Coefficients**	**R^2**
1	PSMCS134	0.62	0.40
2	TFMCS	0.27	0.78
3	PSMI131	0.32	0.86
4	PSMCS137	0.19	0.89
5	TFMI131	0.13	0.90
6	IF1	0.11	0.92
7	AF	- 0.10	0.92
8	TFMSR	0.09	0.93
9	FRCTCMP	0.09	0.94
10	PSMSR90	0.09	0.94

Table II. Overview of Comparison of the Methods

Methods	Applicability	Computational Issues	Representation of Sensitivity	Best Use of Methods
Nominal Range Sensitivity Analysis	Deterministic model.	Need nominal range for each input, potentially time consuming.	Ratios, percentages. Does not include effect of interactions or correlated inputs. Easy to understand.	Key inputs for linear models, verification and validation.
ΔLOR	Deterministic model with output as a probability.	Need nominal range for each input, potentially time consuming.	Ratios, percentages. Does not include effect of interactions or correlated inputs. Easy to understand.	Key inputs for linear models, verification and validation.
Breakeven Analysis	Models used to choose among alternatives.	Complex for model with many decision options and/or more than two inputs, potentially time consuming.	Graphical representation.	Robustness of solution.
AD	Locally different-tiable models.	Require specific software (*e.g.*, ADIFOR)	Local sensitivity measures, such as sensitivity coefficients	Potential key inputs, verification.
Regression	To results from probabilistic simulation	Must specify functional form, computation time and value of solution depends on specific techniques used.	R^2, t-ratios for regression coefficients, standard regression coefficients, and others.	key inputs, joint effect of multiple inputs, verification.
ANOVA (MF)[a]	Probabilistic models	Time consuming for a large number of inputs with interactions.	F-value, Tukey test coefficients, and other that are calculated at different stages of ANOVA.	key inputs, joint effect of multiple inputs, verification.
RSM	Any deterministic model	Developed using a variety of techniques, some require functional forms, others do not; may require extensive runs to generate a calibration data set.	Graphical, evaluation of functional form, method-dependent measures.	Model of models, used with other SA methods to save time.
FAST (MF)	Probabilistic models	Better with no interactions/higher order input. Caution against discrete inputs.	Portion of output variance attributable to each input	Key inputs including combined effect, verification.
MII (MF)	Probabilistic model.	Complex, no computer code available, Time consuming.	Amount of "mutual information" about the output provided my each input, also graphs of intermediate stages.	Key inputs including combined effect.
Scatter Plots (MF)	Probabilistic model.	Easy, time requirement depends on the number of input/outputs.	Graphical, no quantitative sensitivity.	Verification and validation.

[a] MF = denotes a model-independent approach