

A Perspective on the State of the Science of Probabilistic Assessment for Human Exposure

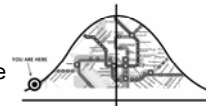
H. Christopher Frey, Ph.D.
Professor

NC STATE UNIVERSITY

Department of Civil, Construction, and Environmental Engineering
North Carolina State University
Raleigh, NC 27695

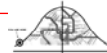
Prepared for:

Society of Toxicology
Probabilistic Risk Assessment (PRA): Bridging
Components Along the Exposure-Dose-Response
Continuum



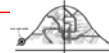
25 – 27 July 2005
Washington DC

NC STATE UNIVERSITY



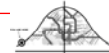
Outline

- Context of Probabilistic Exposure Assessment
- Iteration and Tiers
- Overview of the State of the Science
 - Statistical Methods Based Upon Empirical Data
 - Statistical Methods Based Upon Judgment
 - Other Quantitative Methods
 - Qualitative Methods
 - Sensitivity Analysis
 - Scenario Uncertainty
 - Model Uncertainty
 - Communication
- Conclusions

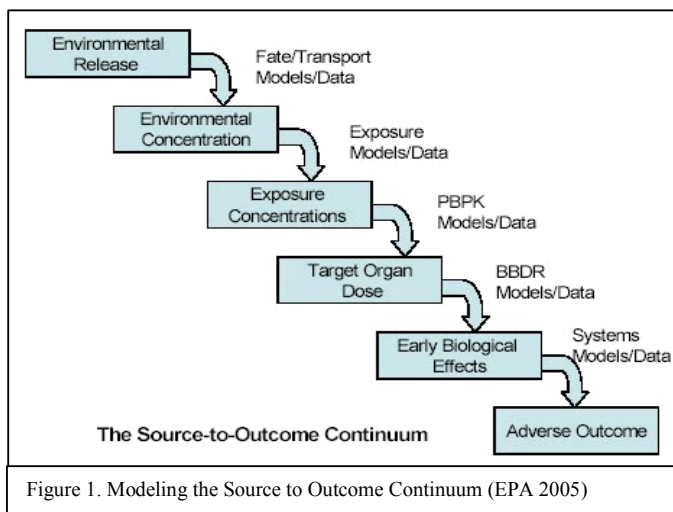


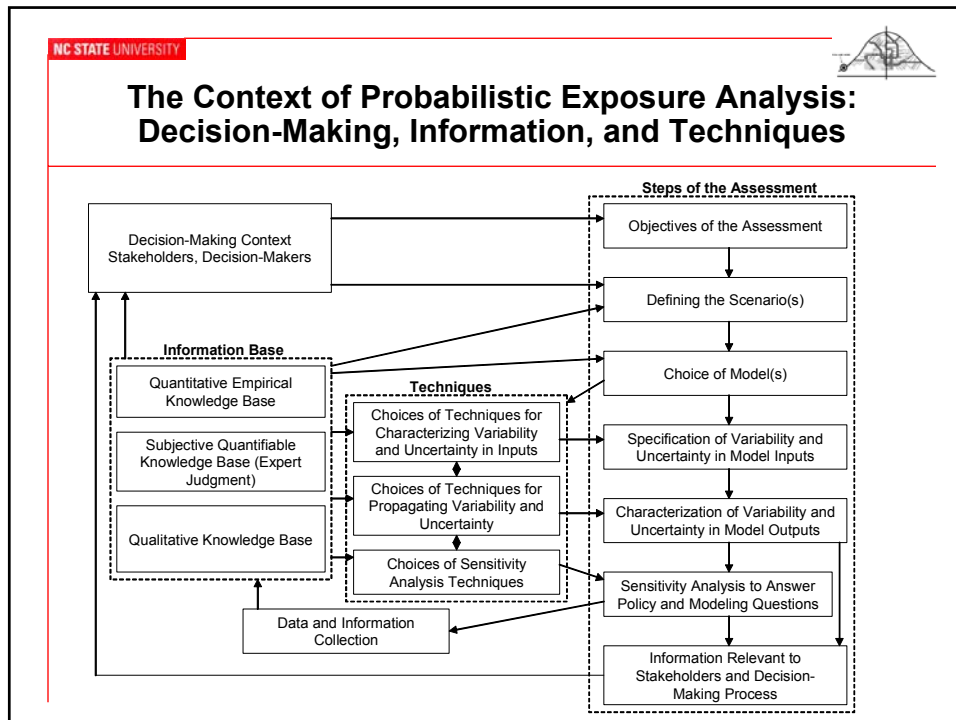
Historical Background on Probabilistic Exposure Assessment

- 1983 “Redbook”
- 1994 “Bluebook”
- US EPA 1997 “Guiding Principles”
- USDA
- FDA
- International
- A current example: WHO/IPCS working group on probabilistic exposure assessment
- Conference presentations
- Peer-reviewed literature

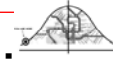


The Context of Exposure Assessment in Risk Analysis



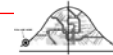


- NC STATE UNIVERSITY
- ### Questions that Decision-Makers and Stakeholders Typically Ask
- How well do we know these numbers?
 - What is the precision of the estimates?
 - Is there a systematic error (bias) in the estimates?
 - Are the estimates based upon measurements, modeling, or expert judgment?
 - How significant are differences between two alternatives?
 - How significant are apparent trends over time?
 - How effective are proposed control or management strategies?
 - What is the key source of uncertainty in these numbers?
 - How can uncertainty be reduced?



Implications of Uncertainty in Decision Making

- Risk preference
 - Risk averse
 - Risk neutral
 - Risk seeking
- Utility theory
- Benefits of quantifying uncertainty: Expected Value of Including Uncertainty
- Benefits of reducing uncertainty: Expected Value of Perfect Information



Some Key Issues

- Clearly define assessment end-point and objectives
- Iterate on model building and input distributions, in combination with sensitivity analysis, to focus resources on the most important inputs/parts of the model
- Conditional probabilistic analyses – mixing point estimates and distributions
- Development of input distributions:
 - Variability
 - Uncertainty
 - Both



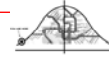
Key Ongoing Challenges

- Superficially, data may seem relevant
- Often, they are still surrogate data
- Lack of directly relevant data
 - Compounds of interest
 - Subpopulations
 - Geographic extent
 - Averaging time
 - Seasonality
 - Etc.



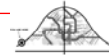
Key Ongoing Challenges

- Mis-specification of uncertainty in 2D analyses:
 - Arbitrarily specifying “uncertainty” in distribution parameters and randomly sampling from them could lead to nonsense
 - Correlation/dependencies in sampling distributions of distribution parameters can be complex
 - Can obtain wildly inaccurate or meaningless results if this issue is not carefully addressed

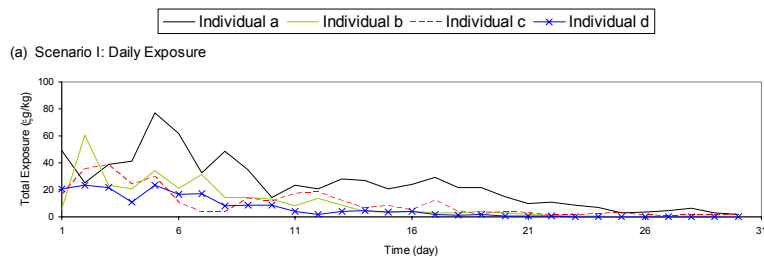


Key Ongoing Challenges

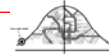
- Expert Judgment vs. Data
 - Perception that judgment is more biased than analysis of available data
 - Unless data are exactly representative, they too could be biased
 - Statistical methods are “objective” in that the results can be reproduced by others, but this does not guarantee absence of bias
 - A key area for moving forward is to agree on conditions under which expert judgment is an acceptable input, even for rulemaking situations



Key Ongoing Challenges: Time and Space

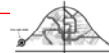


- Exposures vary over time for each individual
- Individual activity occurs over a geographic area
- Implications for probabilistic analysis – e.g., inter-individual vs. intra-individual dependencies, sensitivity analysis?



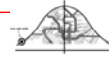
Key Ongoing Challenges

- The EPA “Guiding Principles” is a useful starting point.
- “Model uncertainty” and “scenario uncertainty” deserve explicit guiding principles
- Other needs:
 - » Communication of probabilistic information
 - » Use of probabilistic information in decision making



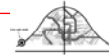
Variability and Uncertainty

- **Variability:** refers to the certainty that
 - different members of a population will have different values (inter-individual variability)
 - values will vary over time for a given member of the population (intra-individual variability)
- **Uncertainty:** refers to lack of knowledge regarding
 - True value of a fixed but unknown quantity
 - True population distribution for variability
- Both depend on averaging time



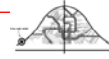
Variability and Uncertainty

- Sources/Characteristics of Variability
 - Inherent differences, such as because of natural or behavioral factors
 - Stochasticity – inherent differences that cannot currently be explained
 - Periodicity, seasonality
 - Mixtures of subpopulations
 - Variation that could be explained with better models
 - Variation that could be reduced through control measures



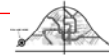
Variability and Uncertainty

- Sources of Uncertainty:
 - Random sampling error for a random sample of data
 - Measurement errors
 - » Systematic error (bias, lack of accuracy)
 - » Random error (imprecision)
 - Non-representativeness
 - » Not a random sample, leading to bias in mean (e.g., only measured loads not typical of daily operations)
 - » Direct monitoring versus infrequent sampling versus estimation, averaging time
 - » Omissions
 - Surrogate data (analogies with similar sources)
 - Lack of relevant data
 - Problem and scenario specification
 - Modeling



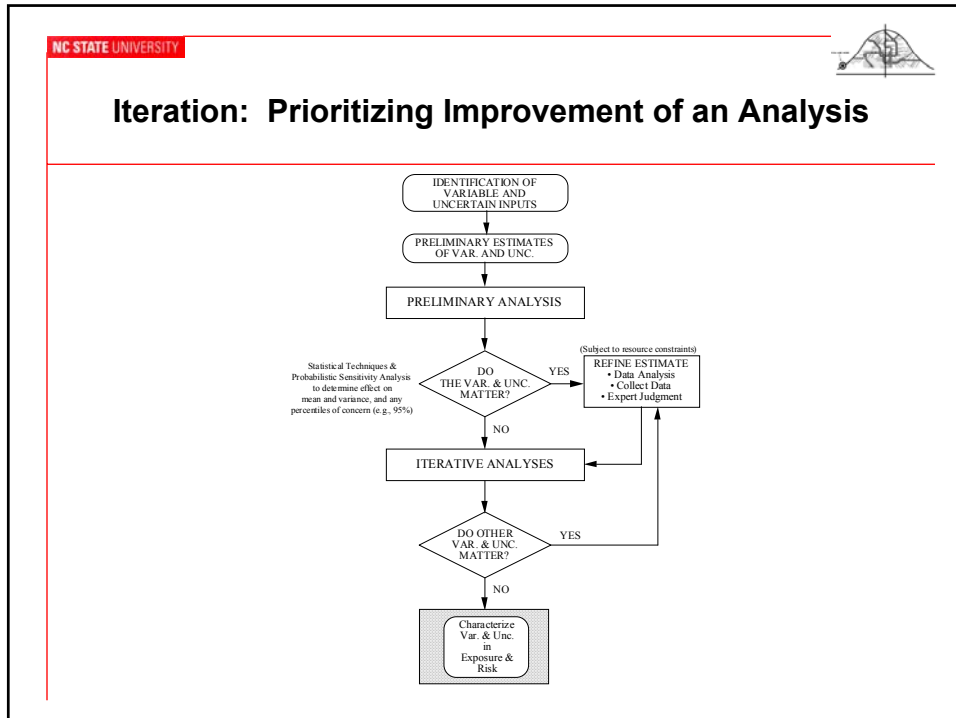
Tiered Approach to Analysis

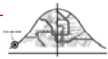
- Purpose of Analyses (examples)
 - Screening to prioritize problems
 - Regulatory decision-making
 - Research planning
- Objectives of an Analysis
 - Individual, population or both
 - Acute, long-term, or both?
- Data quality objectives: how good do the estimates need to be, or can they be?

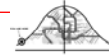


Tiers of Analysis: Examples

- Maximally Exposed Individual (MEI)
- Theoretical Upper Bound Estimate (TUBE)
- High-End and Central Tendency Exposure
- “One-Dimensional” simulation of variability
- “One-Dimensional” simulation of uncertainty
- “Two-Dimensional” simulation of variability and uncertainty

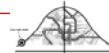


- NC STATE UNIVERSITY
- 
- ### Overview of “State of the Science”
- Scenario Uncertainty
 - Model Uncertainty
 - Statistical Methods Based Upon Empirical Data
 - Statistical Methods Based Upon Judgment
 - Other Quantitative Methods
 - Qualitative Methods
 - Sensitivity Analysis
 - Communication



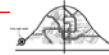
Scenario Uncertainty

- Key dimensions
 - Pollutants
 - Transport pathways
 - Exposure routes
 - Susceptible populations
 - Averaging time
 - Geographic extent
 - Time Periods
 - Activity Patterns
- Which dimensions/combinations matter, which ones don't?
- Uncertainty associated with mis-specification of a scenario – systematic error
 - descriptive errors
 - aggregation errors
 - errors in professional judgment
 - incomplete specification of the scenario



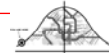
Model Uncertainty

- Conceptual model vs. reality
- Simplifications
 - Aggregation
 - Exclusion
- Structure
- Level of detail
- Resolution
- Model Boundaries (related to scenario)
- Boundary conditions
- Calibration
- Validation, Partial validation
- Extrapolation
- Precision and accuracy



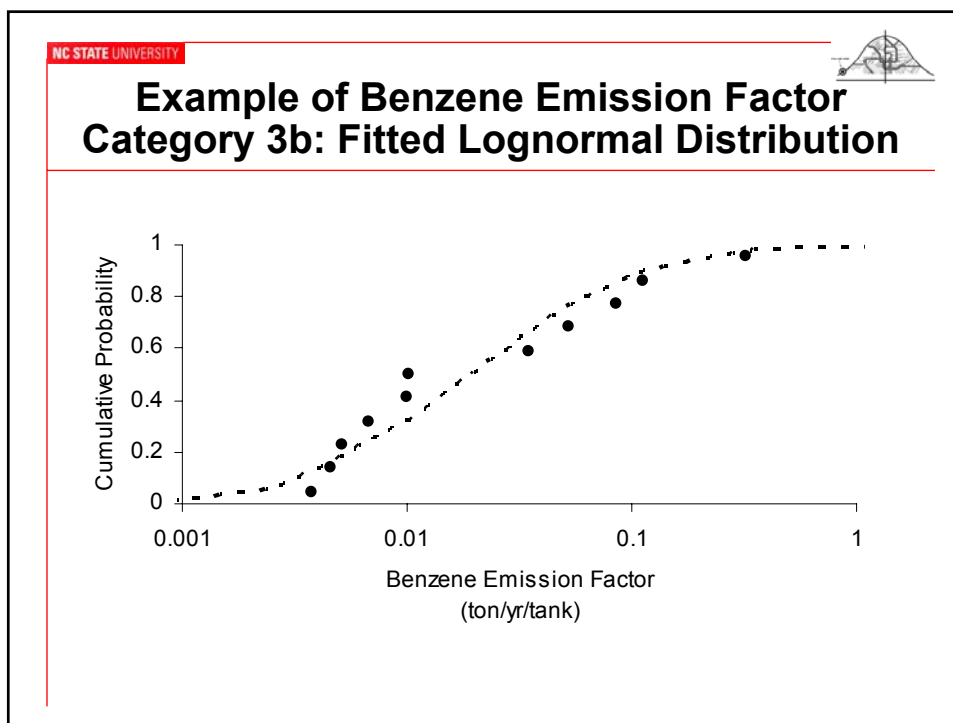
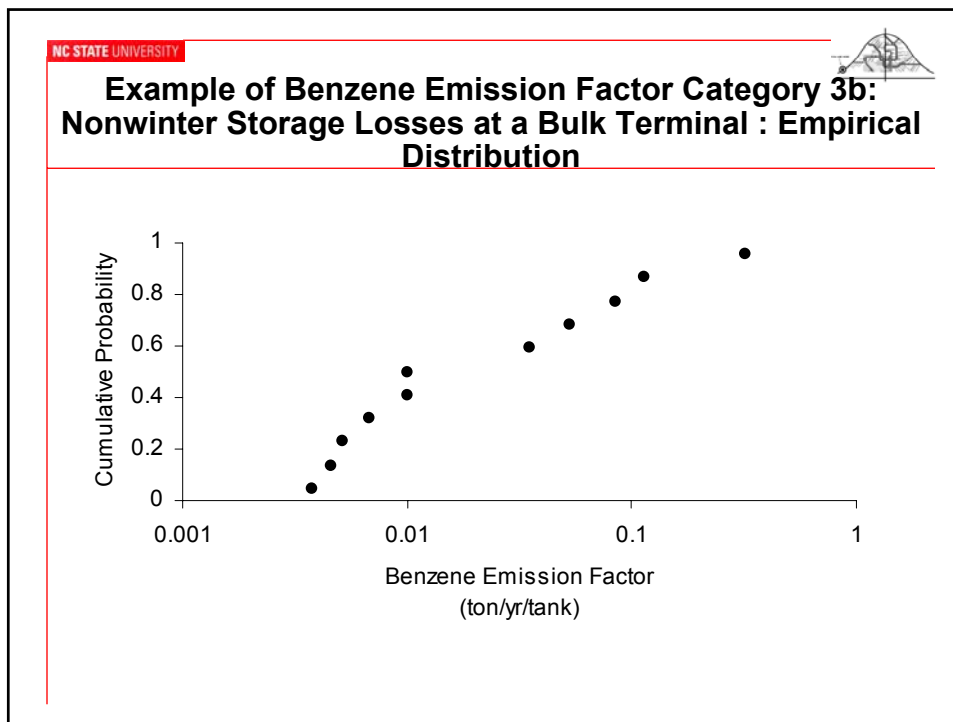
Model Uncertainty

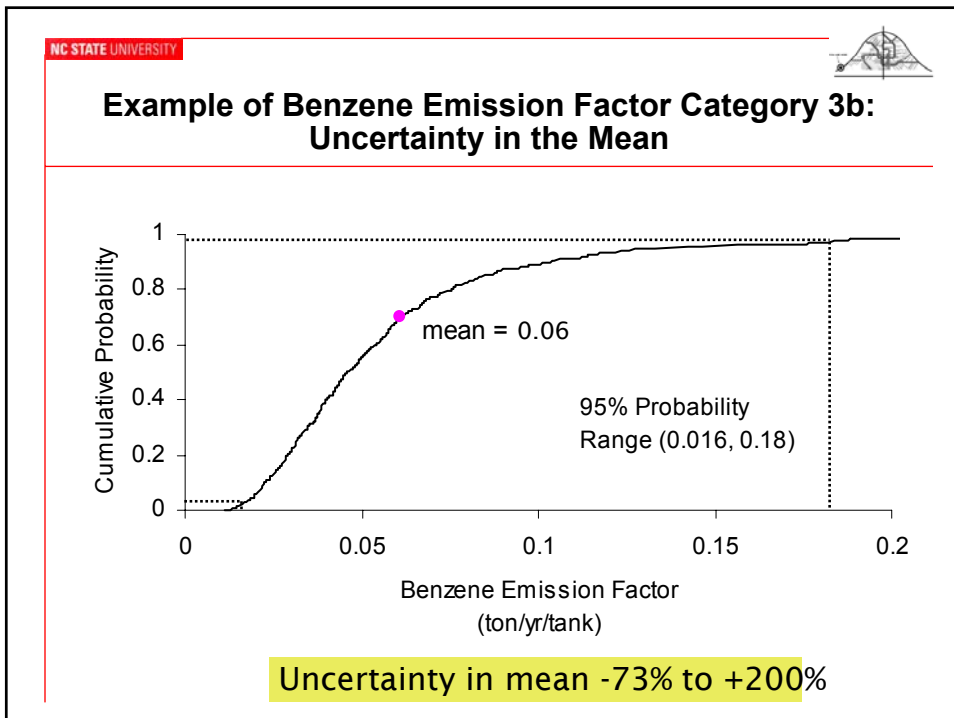
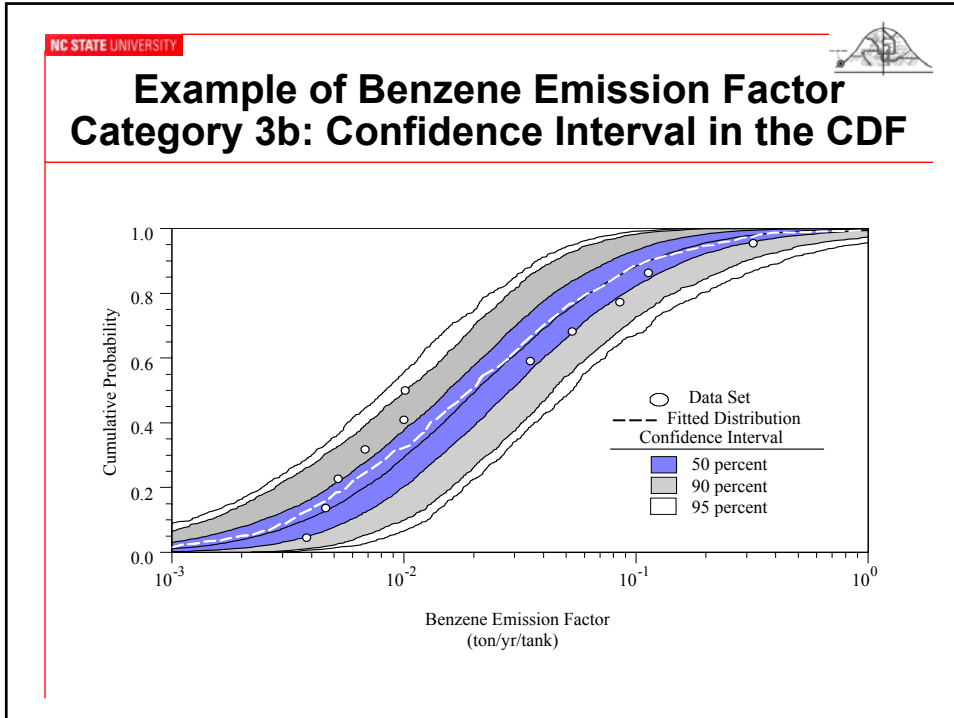
- Methods for Dealing with Model Uncertainty
 - Compare alternative models, but do not combine
 - Weight predictions of alternative models (e.g., probability trees)
 - Meta-models that degenerate into alternative models

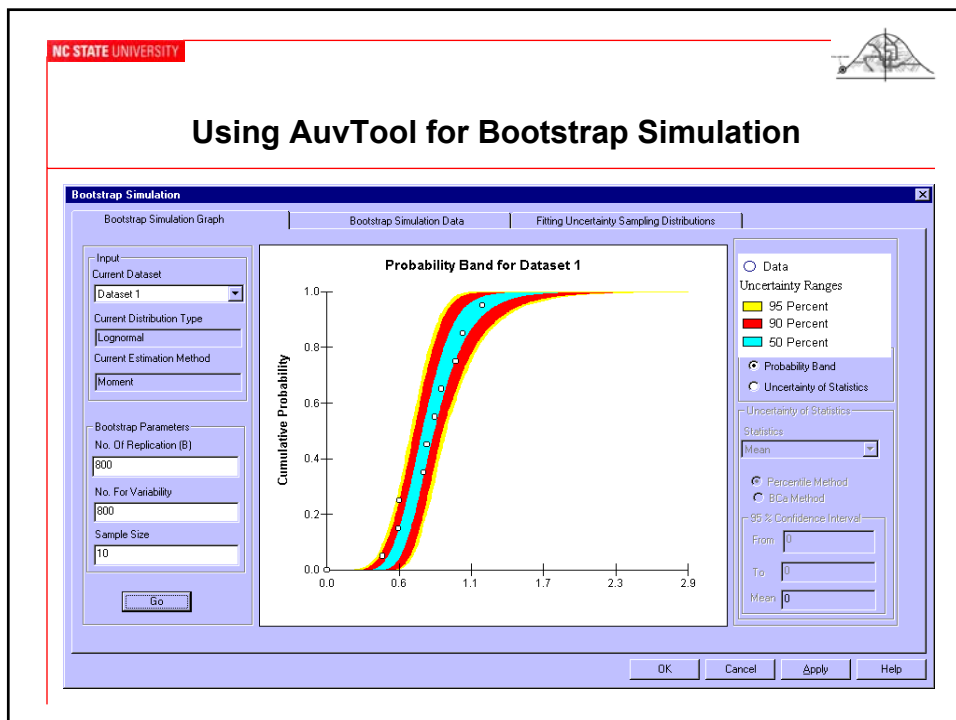
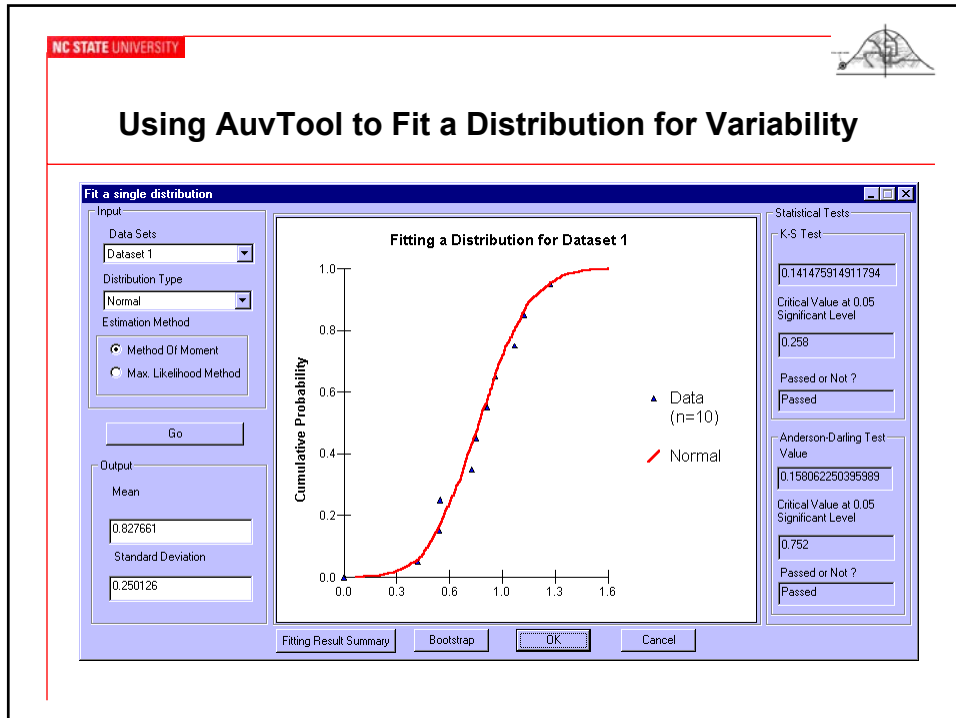


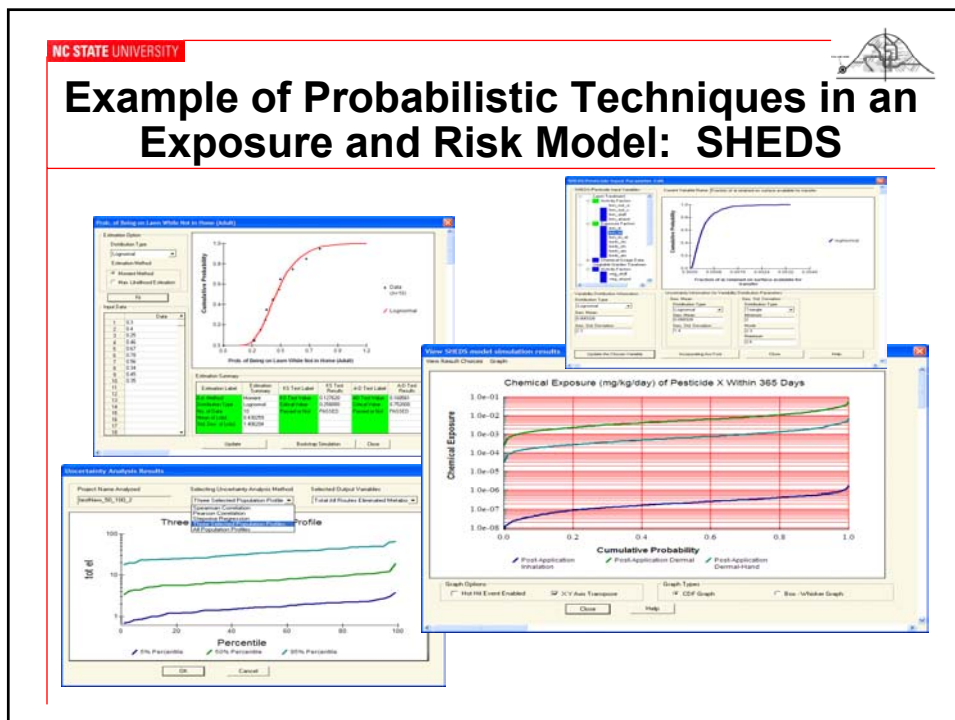
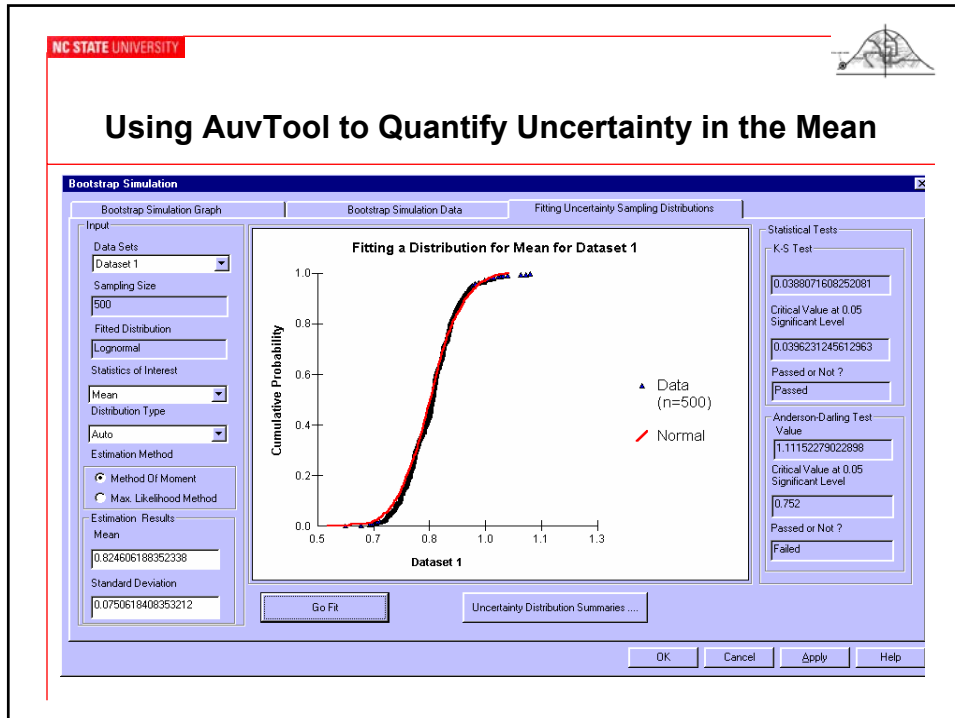
Statistical Methods Based Upon Empirical Data

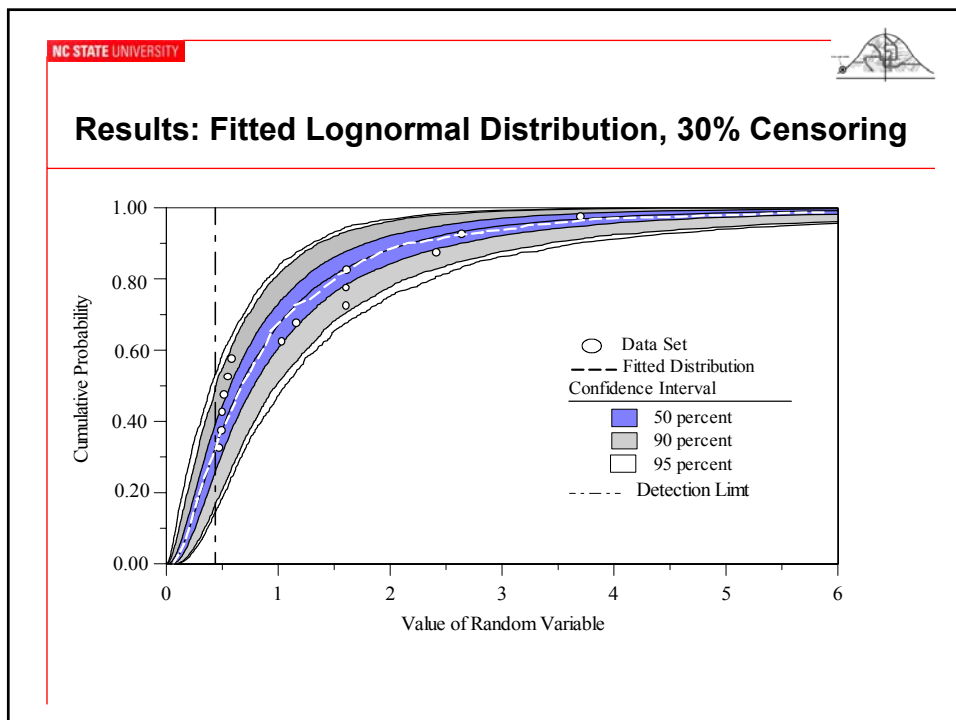
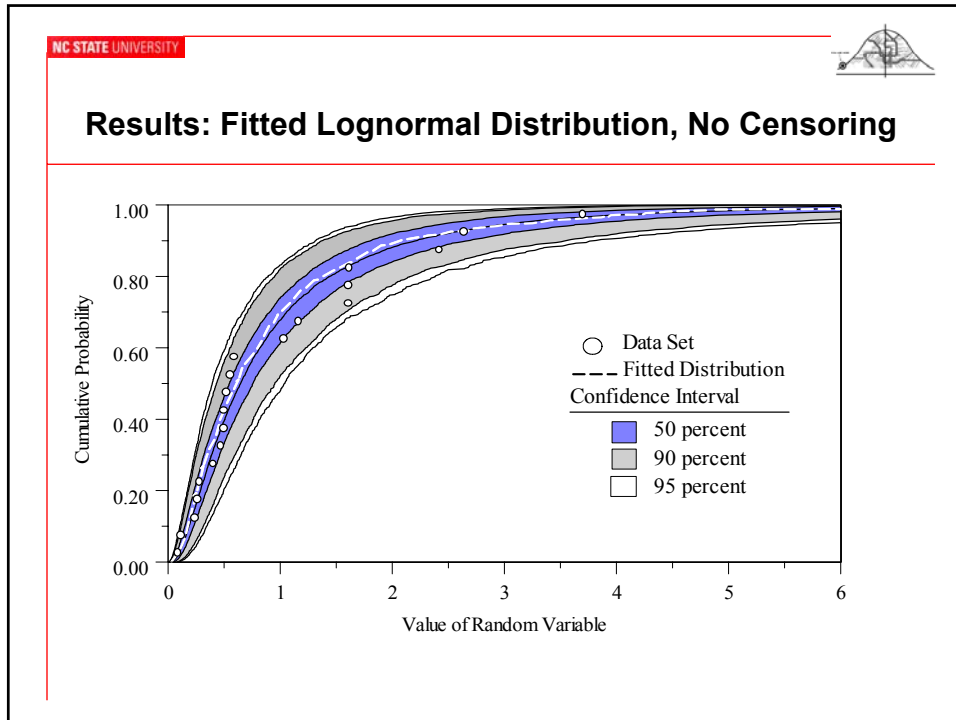
- Frequentist, classical
- Statistical inference from sample data
 - Parametric approaches
 - » Parameter estimation
 - » Goodness-of-fit
 - Nonparametric approaches
 - Mixture distributions
 - Censored data
 - Dependencies, correlations, deconvolution
 - Time series, autocorrelation

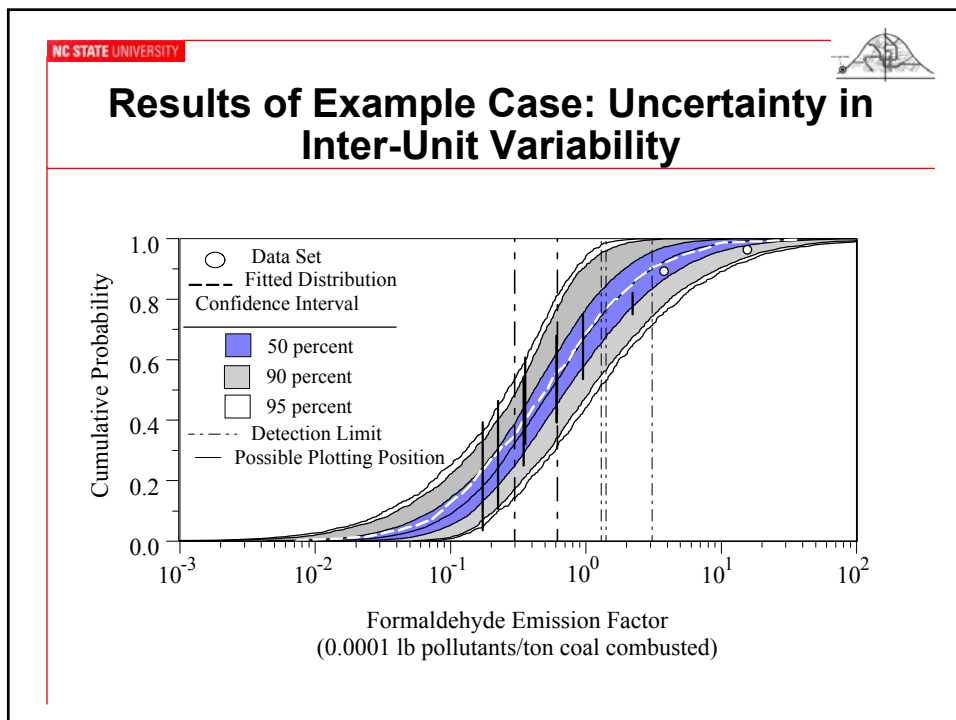
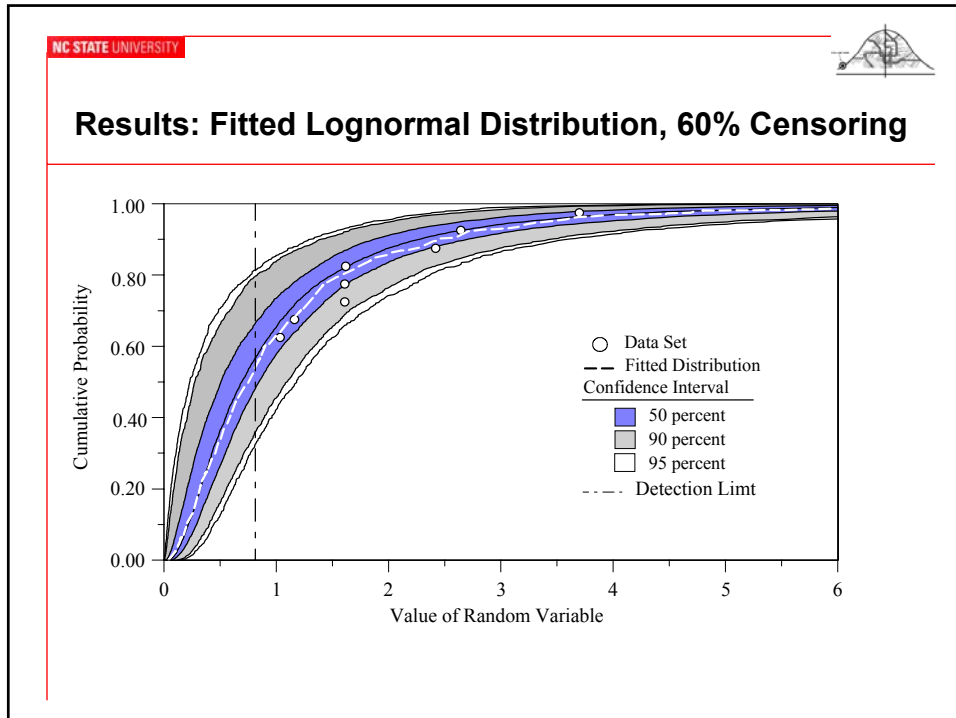


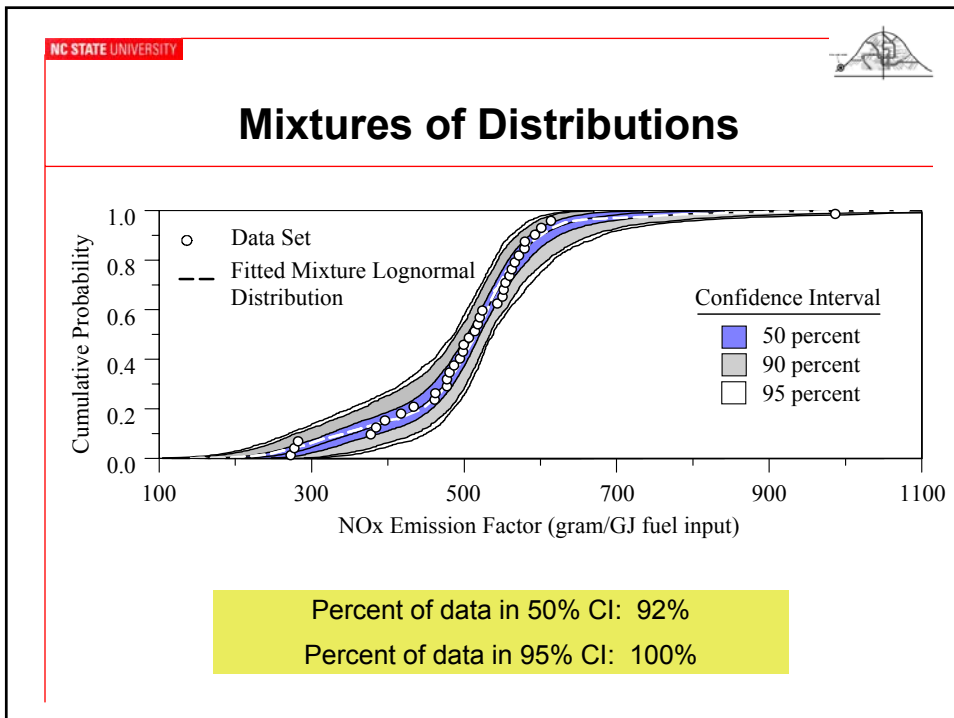
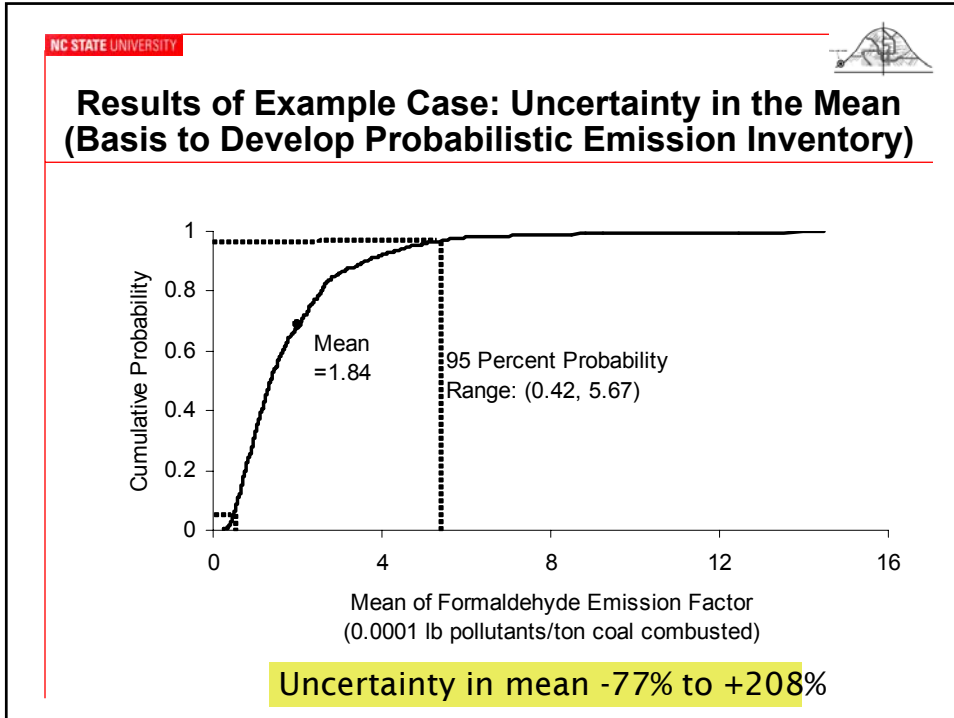


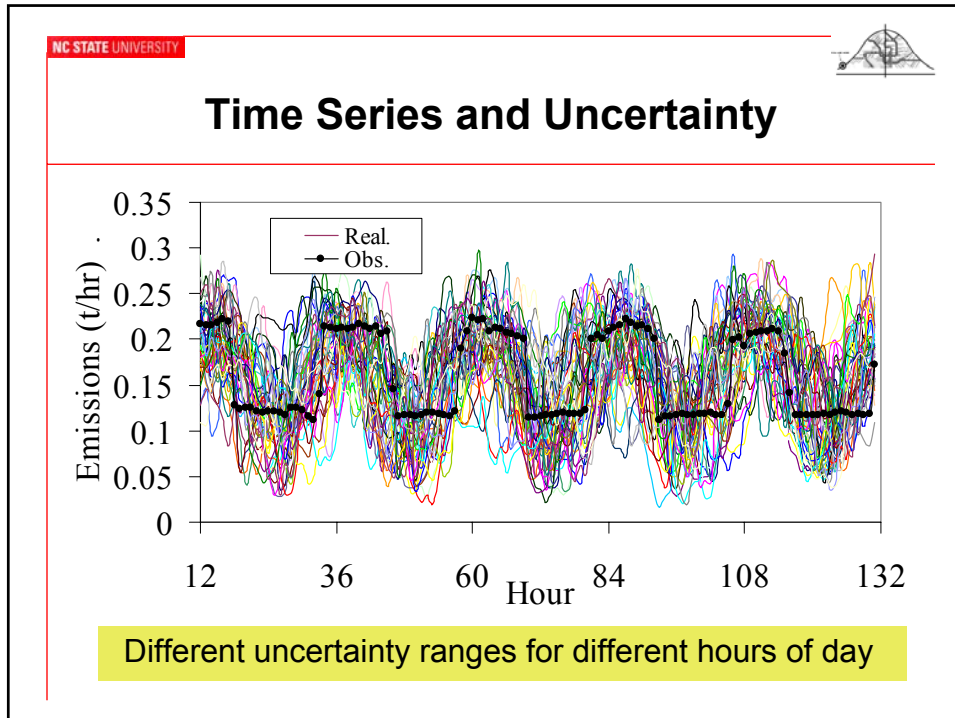


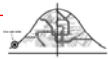









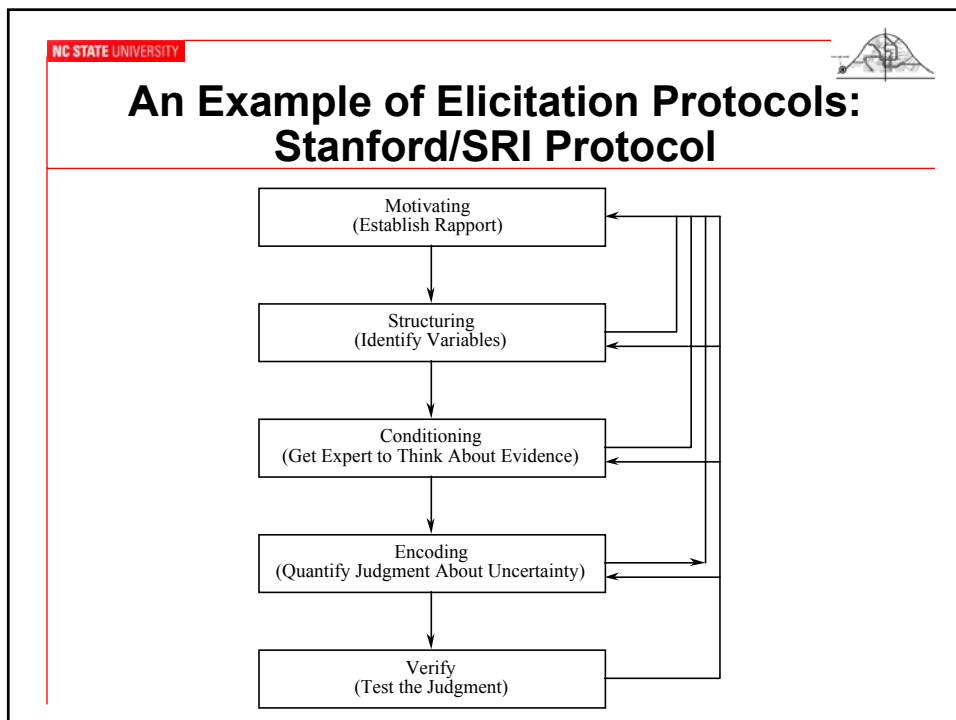


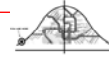
- NC STATE UNIVERSITY
- 
- ## Propagating Variability and Uncertainty
- Analytical techniques
 - » Exact solutions (limited applicability)
 - » Approximate solutions
 - Numerical methods
 - » Monte Carlo
 - » Latin Hypercube Sampling
 - » Other sampling methods (e.g., Hammersley, Importance, stochastic response surface method, Fourier Amplitude Sensitivity Test, Quasi-Monte Carlo methods, etc.)

NC STATE UNIVERSITY 

Statistical Methods Based Upon Expert Judgment

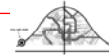
- Expert Elicitation
 - Heuristics and Biases
 - » Availability
 - » Anchoring and Adjustment
 - » Representativeness
 - » Others (e.g., Motivational, Expert, etc.)
 - Elicitation Protocols
 - » Motivating the expert
 - » Structuring
 - » Conditioning
 - » Encoding
 - » Verification
 - Documentation
 - Individuals and Groups
 - When Experts Disagree





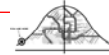
Statistical Methods Based Upon Expert Judgment

- Bayesian methods can incorporate expert judgment
 - Prior distribution
 - Update with data using likelihood function and Bayes' Theorem
 - Create a posterior distribution
- Bayesian methods can also deal with various complex situations:
 - Conditional probabilities (dependencies)
 - Combining information from multiple sources
- Appears to be very flexible
- Computationally, can be very complex
- Complexity is a barrier to more widespread use



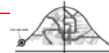
Other Quantitative Methods

- Interval Methods
 - Simple intervals
 - Probability bounds
 - Produce “optimally” narrow bounds – cannot be any narrower and still enclose all possible outcomes, including dependencies among inputs
 - Bounds can be very wide in comparison to confidence intervals




Other Quantitative Methods

- Fuzzy methods
 - Representation of vagueness, rather than uncertainty
 - Approximate/semi-quantitative
 - Has been applied in many fields
- Meta-analysis
 - Quantitatively combine, synthesize, and summarize data and results from different sources
 - Requires assessment of homogeneity among studies prior to combining
 - Produces data with larger sample sizes than the constituent inputs
 - Can be applied to summary data
 - If raw data are available, other methods may be preferred




Qualitative Methods

- Principles of rationality
- Lines of reasoning
 - Direct empirical evidence (direct measurement)
 - Semi-empirical evidence (surrogate data)
 - Empirical correlations (more distant surrogates)
 - Theory-based inference
 - Existential insight (expert judgment)
- Weight of evidence
 - Proof on one side is greater than proof on another side
 - Consider “evidence factors”

NC STATE UNIVERSITY 

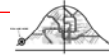
Sensitivity Analysis

- Objectives of Sensitivity Analysis (examples):
 - Help identify key sources of variability (to aid risk management strategy)
 - Help identify key sources of uncertainty (to prioritize additional data collection to reduce uncertainty)
 - Variance of an output
 - What causes worst/best outcomes
 - What are critical control points, critical limits
- Local vs. Global Sensitivity Analysis
- Model Dependent vs. Model Independent Sensitivity Analysis
- Applicability of methods often depends upon characteristics of a model (e.g., nonlinear, thresholds, categorical inputs, etc.)

NC STATE UNIVERSITY 

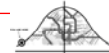
Examples of Sensitivity Analysis Methods

<ul style="list-style-type: none">• Mathematical Methods → Assess sensitivity of a model output to the range of variation of an input.	Nominal Range Sensitivity Analysis (NRSA)
	Differential Sensitivity Analysis
<ul style="list-style-type: none">• Statistical Methods → Effect of variance in inputs on the output distribution.	Regression Analysis (RA)
	Analysis of Variance (ANOVA)
	Classification and Regression Trees (CART)
<ul style="list-style-type: none">• Graphical Methods → Representation of sensitivity in the form of graphs, charts, or surfaces.	Scatter Plots
	Conditional Sensitivity



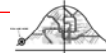
Sensitivity Analysis Methods (Examples)

- Nominal Range Sensitivity Analysis
- Differential Sensitivity Analysis
- Conditional Analysis
- Correlation coefficients (sample, rank)
- Linear regression (sample, rank, variety of basis functions possible)
- Other regression methods
- Analysis of Variance (ANOVA)
- Categorical and Regression Trees (CART) (a.k.a. Hierarchical Tree-Based Regression)
- Sobol's method
- Fourier Amplitude Sensitivity Test (FAST)
- Mutual Information Index



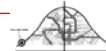
Sensitivity Analysis: Displays/Summaries

- Scatter plots
- Line plots/conditional analyses
- Radar plots
- Distributions (for uncertainty or variability in sensitivity)
- Summary statistics
- Categorical and regression trees
- Apportionment of variance

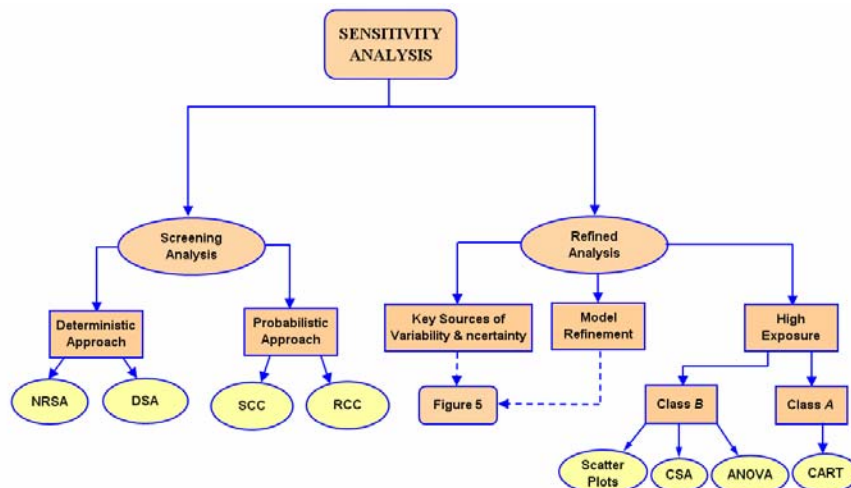


Guidance on Sensitivity Analysis

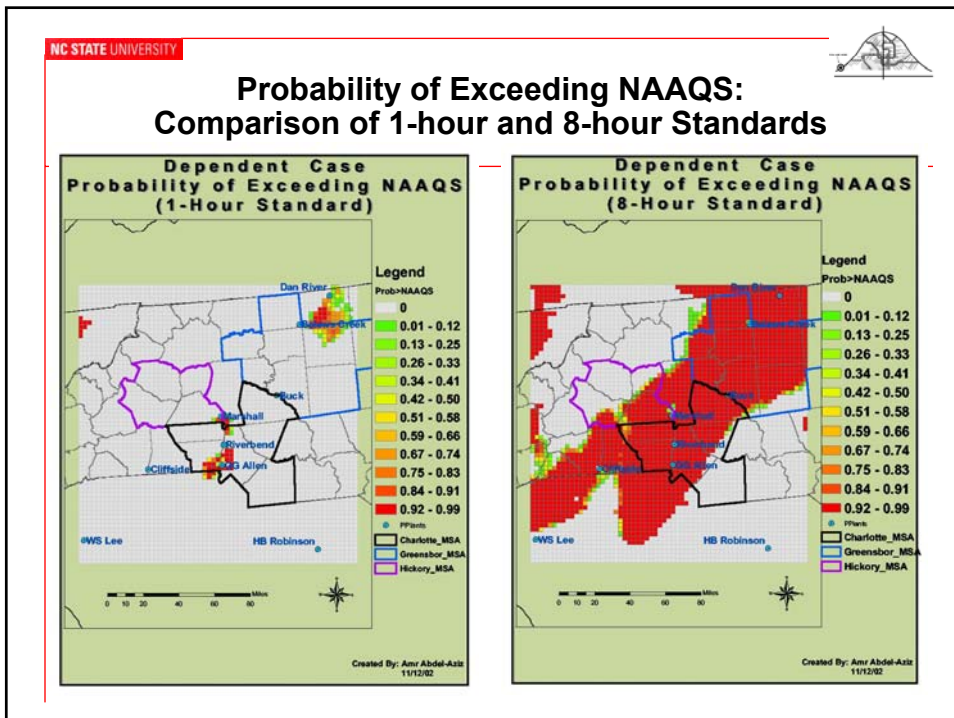
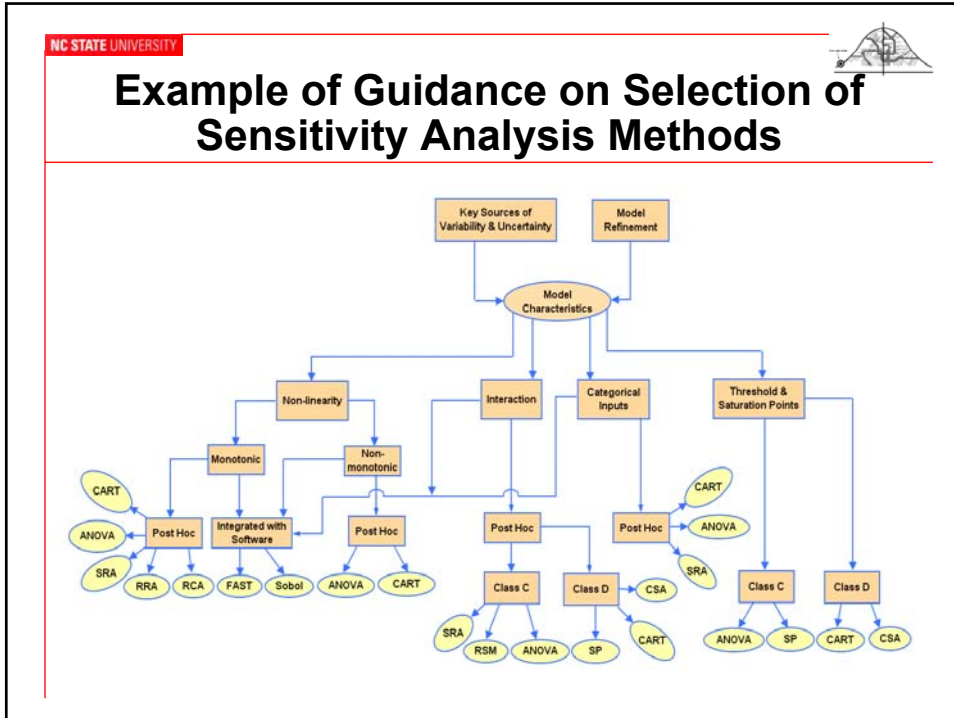
- When to perform sensitivity analysis
- Information needed depending upon objectives
- Preparation of existing or new models
- Defining the case study/scenarios
- Selection of sensitivity analysis methods
- Procedures for application of methods
- Presentation and interpretation of results

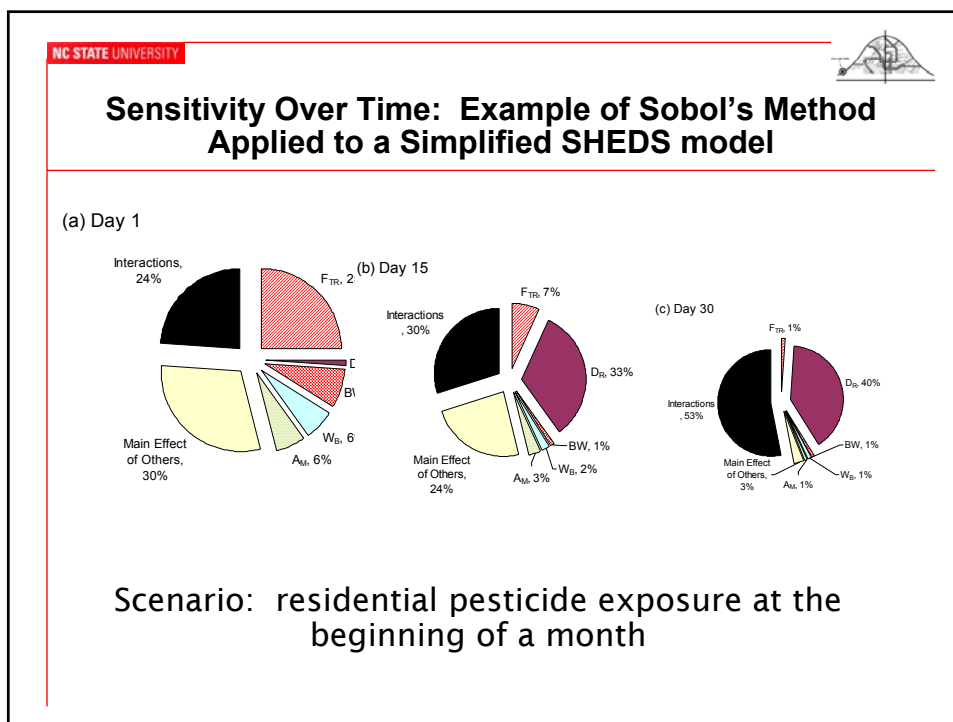
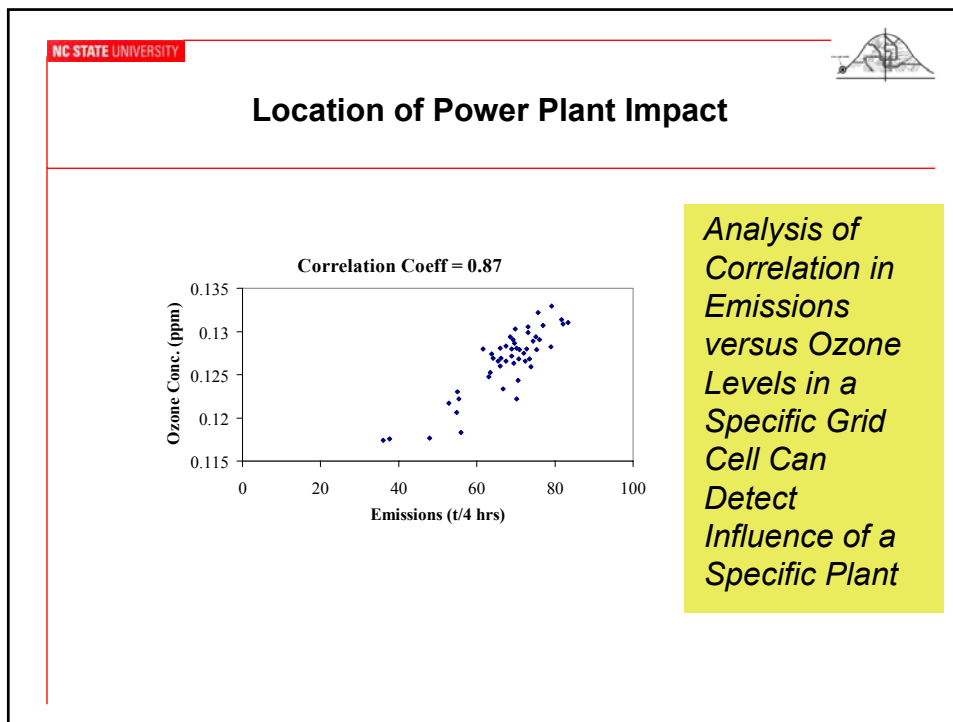


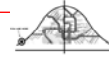
Example of Guidance on Selection of Sensitivity Analysis Methods



Source: Frey et al., 2004, www.ce.ncsu.edu/risk/

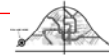






Communication

- Case Studies (scenarios)
- Graphical Methods
 - Influence Diagrams
 - Decision Tree
 - Others
- Summary statistics/data
- Evaluation of effectiveness of methods for communication (e.g., Bloom et al., 1993; Ibrekk and Morgan, 1987)



Summary

- Probabilistic Exposure Assessment has become more common in the last 10 years
- Both empirical and judgment based approaches are used, but in fact all approaches require judgment
- Monte Carlo and similar methods are widely used because of their flexibility
- Areas of growing need or recognition include sensitivity analysis, model uncertainty, scenario uncertainty, communication, and linkages to decision-making
- Opportunities remain for probabilistic methods to improve (e.g., dealing with data quality limitations, growing interest in higher resolution of analysis)