Quantification of Uncertainty in Air Pollutant Emissions Inventories

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:

Emission Inventory Conference
Raleigh, NC

May 16, 2007

Outline

- Significance of Emissions Inventories and Their Uncertainties
- Methods for Uncertainty and Sensitivity Analysis
- Probabilistic Emissions Factors and Inventories
- Findings
- Recommendations

Reasons for Needing Emission Inventories

- Implementation Plan or Control Strategy Development
- Emission Cap and Trade Activities
- Early Reduction Program Design
- Emission Trends Analysis and Projections
- Permit Limit Determination
- Information for Public
- Emission Statement/Fee Collection
- International Treaty Reporting
- Environmental Impact Modeling and Assessment
- Field Study Design
- Compliance Determination
- Real-time Air Quality Forecasting
- Conformity Analysis
- Exposure and Risk Analysis
- Prioritizing Data Needs
- Accountability Assessments

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 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?

Recommendations for Quantification of Uncertainty in Emissions

- National Research Council
  - 1992, *Rethinking the Ozone Problem in Urban and Regional Air Pollution*
  - 1994, *Science and Judgment in Risk Assessment*
  - 2000, *Modeling Mobile Source Emissions*
- NARSTO Emission Inventory Assessment (2005)
- Intergovernmental Panel on Climate Change
  - 2000, *Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories*
  - *2006 IPCC Guidelines for National Greenhouse Gas Inventories*
- EPA Office of the Inspector General (2006)

Trends in Quantity and Quality of Emission Factors

Factor Rating	Qualitative Description	As of March 1996		As of September 2004	
		Number of Factors	Percent of Total	Number of Factors	Percent of Total
A	Excellent	1,270	14	2,135	12
B	Above Average	1,190	13	1,829	11
C	Average	1,513	17	2,619	15
D	Below Average	2,077	24	4,740	28
E	Poor	2,788	32	5,787	35
Total		**8,838**		**17,110**	

Source: EPA Office of the Inspector General (2006)

Current Practice for *Qualifying* Uncertainty in Emission Factors and Inventories

- Qualitative ratings for emission factors (AP-42)
- Data Attribute Rating System (DARS) (not really used in practice)
- Both methods are qualitative
- No quantitative interpretation
- Some sources of uncertainty (i.e. non-representativeness) difficult to quantify
- Qualitative methods can complement quantitative methods

Variability and Uncertainty

- **Variability**: refers to the certainty that
  - different emission sources will have different emissions (inter-unit variability)
  - emissions will vary over time for a given source (intra-unit 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

Sources of Variability

- Design
- Feedstocks
- Ambient Conditions
- Maintenance Practices
- Operational practices and occurrences (e.g., load following, baseload, transients, process upsets)
- Seasonality/Periodicity

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

How Good are the Emission Factors? NO_x From Coal-Fired Power Plants

- *Compilation of Air Pollutant Emission Factors*, AP-42, U.S. Environmental Protection Agency
- Example: Bituminous coal, Wall-fired, dry bottom boilers
  - Data Quality Rating of A
  - 28 data points
  - 9.5 to 44.5 lb NO_x/ton
  - Average = 21.1 lb NO_x/ton, std dev = 8.2
  - 95% Confidence Interval on the mean: 17.9 - 24.1 (±15%)
- Uncertainty: Depends on context
  - ±15% of the average of many plants
  - A factor of two for a randomly selected individual plant
- The Rating of an AP-42 emission factor does not translate into a quantitative estimate of uncertainty

Methods for Quantifying Uncertainty

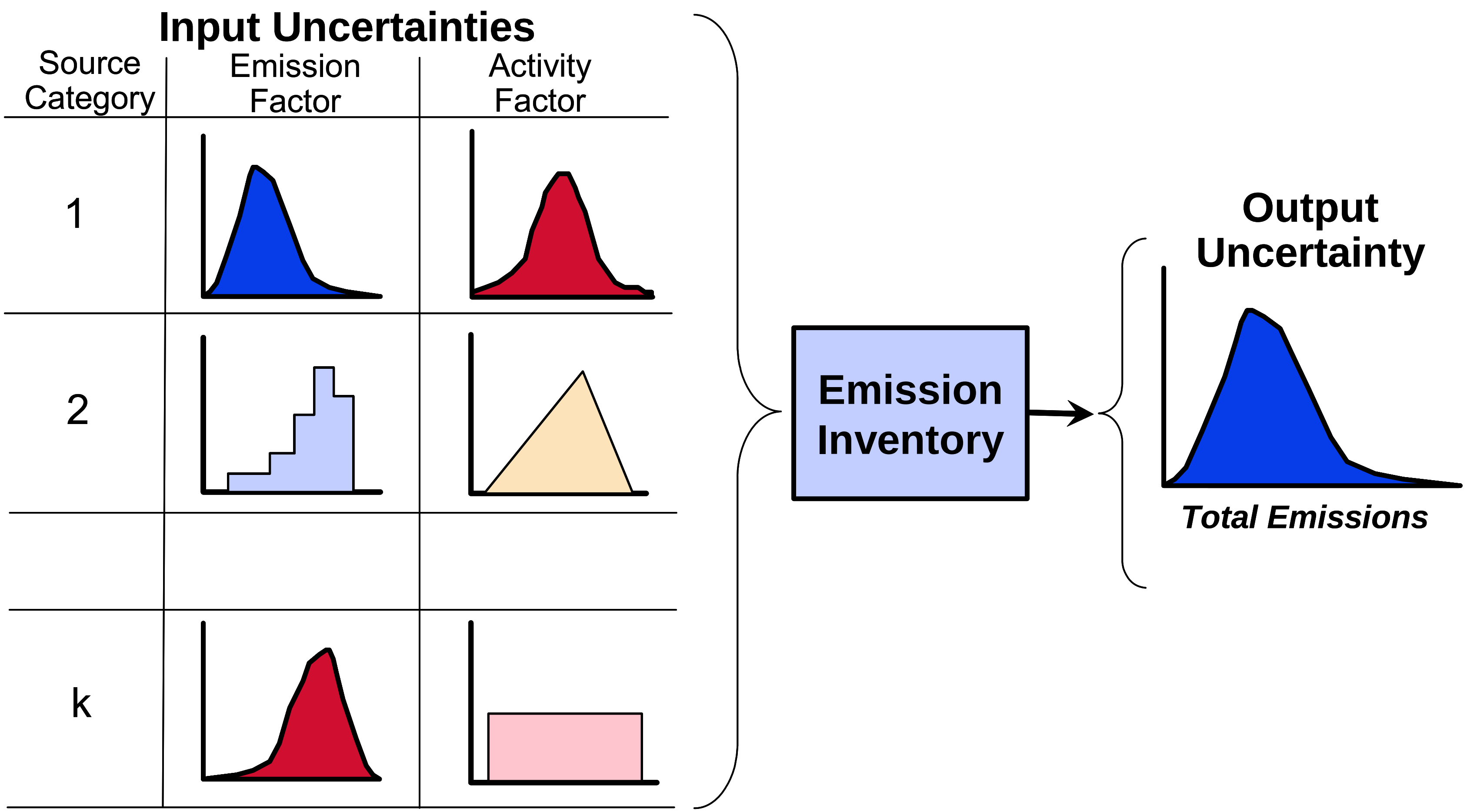
- Bottom-Up Approaches
  - Statistical Methods Based Upon Empirical Data
  - Statistical Methods Based Upon Judgment
  - Sensitivity Analysis
- Top-Down Approaches

Top-Down Methods: Identify Major Biases

- Comparison of Inventories Developed Using Independent Methods and Data
- Comparison of Air Quality Model Predictions to Monitoring Data
- Comparison of Source-Oriented vs. Receptor-Oriented Modeling Approaches
- Comparison of Inventories and Ambient Monitoring Data

NC STATE UNIVERSITY

Quantifying Uncertainties in Emission Inventories: Conceptual Bottom-Up Approach



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

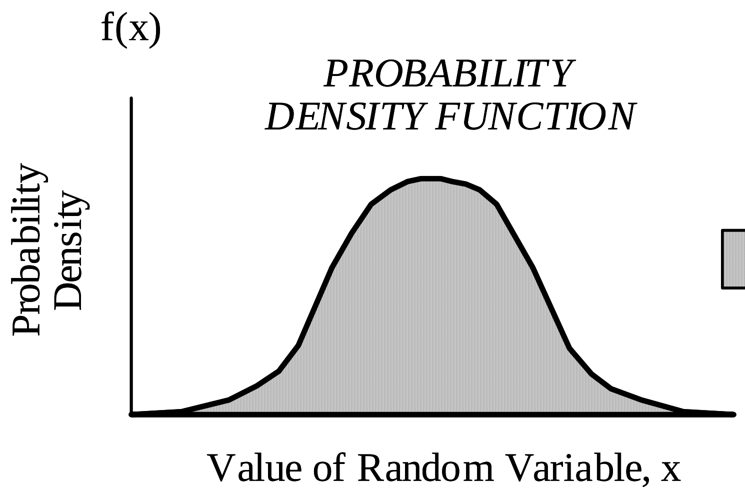
Propagating Variability and Uncertainty

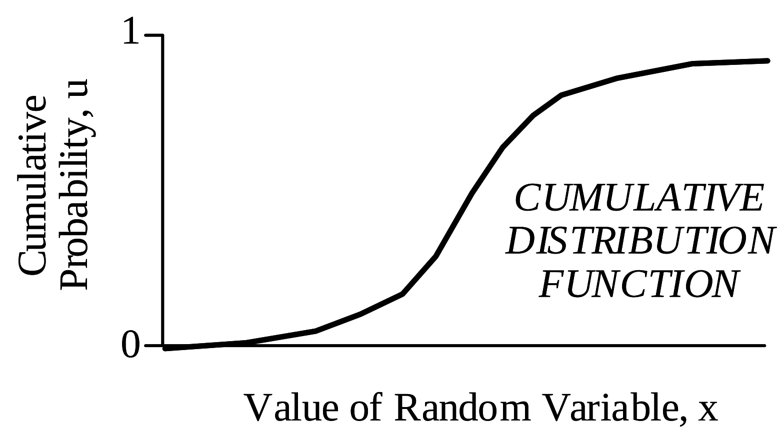
- Analytical techniques
  - Exact solutions (limited applicability)
  - Approximate solutions
- Numerical methods
  - Monte Carlo
  - Latin Hypercube Sampling
  - Other sampling methods (e.g., Hammersley, Importance, stochastic response surface method, Fourier Amplitude Sensitivity Test, Sobol's method, Quasi-Monte Carlo methods, etc.)

Monte Carlo Simulation

- Probabilistic approaches are widely used
- Monte Carlo (and similar types of) simulation are widely used.
- Why?
  - Extremely flexible
    - » Inputs
    - » Models
  - Relatively straightforward to conceptualize

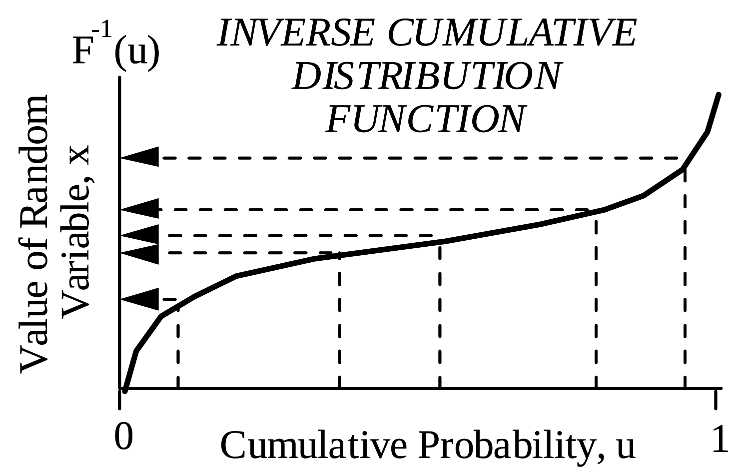
Propagating Distributions Through Models





MONTE CARLO SIMULATION
- Generate a random number u~U(0,1)
- Calculate $F^{-1}(u)$ for each value of u

LATIN HYPERCUBE SAMPLING
- Divide u into N equal intervals
- Select median of each interval
- Calculate $F^{-1}(u)$ for each interval
- Rank each sample based on U(0,1) (or restricted pairing technique)



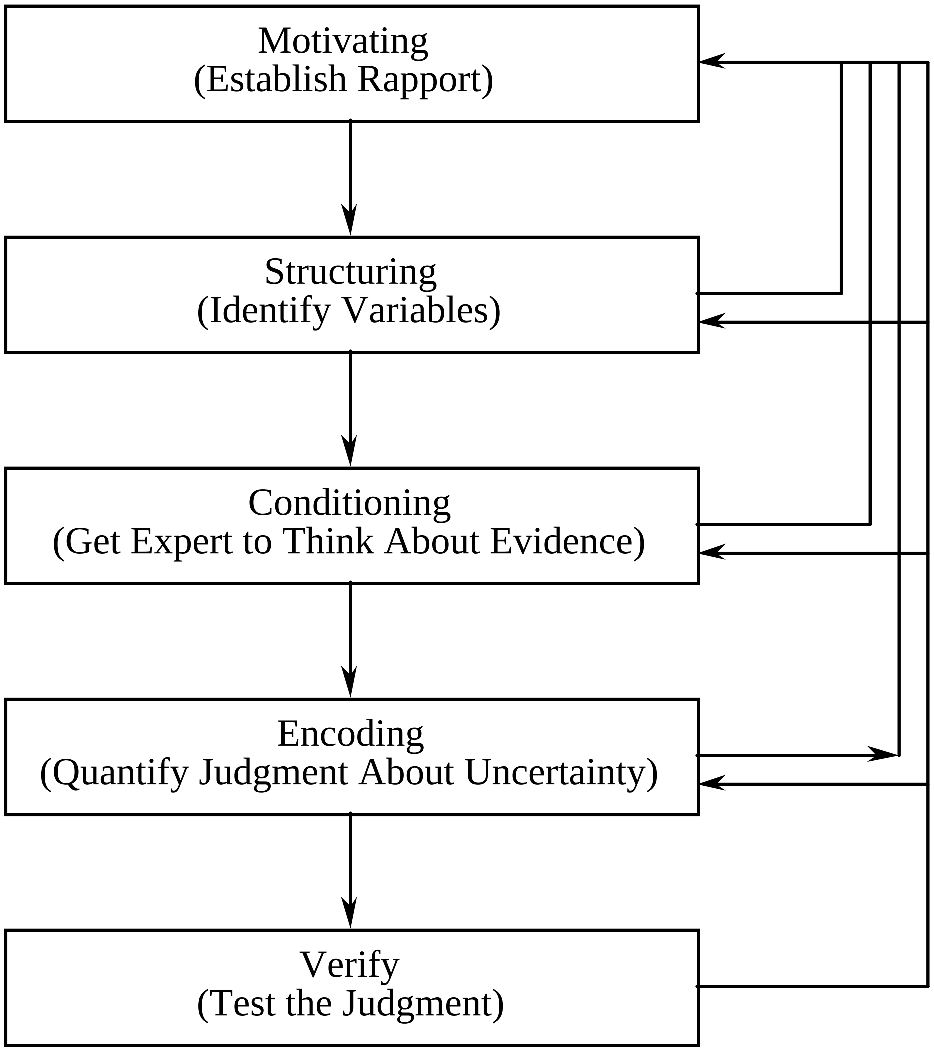
Tiered Approach to Analysis

- Purpose of Analyses (examples)
  - Screening to prioritize resources
  - Regulatory decision-making
  - Research planning
- Types of Analyses
  - Screening level point-estimates
  - Sensitivity Analysis
  - One-Dimensional Probabilistic Analysis
  - Two-Dimensional Probabilistic Analysis
  - Non-probabilistic approaches

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 Diasagree

An Example of Elicitation Protocols: Stanford/SRI Protocol



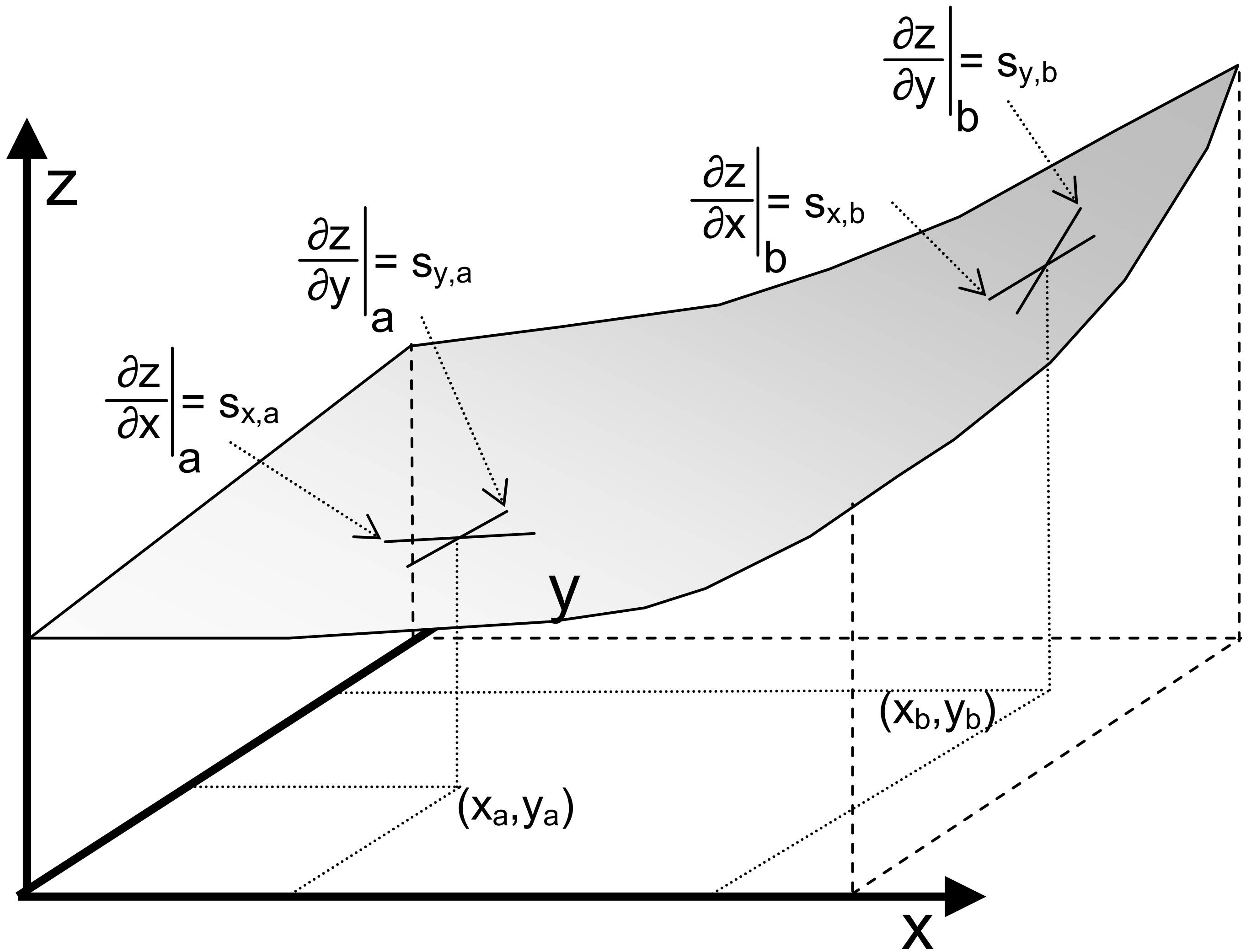
Statistical Methods Based Upon Expert Judgment

- 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

Sensitivity Analysis

- Objectives of Sensitivity Analysis (examples):
  - Help identify key sources of variability (to aid management strategy)
    - Critical control points?
    - Critical limits?
  - Help identify key sources of uncertainty (to prioritize additional data collection to reduce uncertainty)
  - What causes worst/best outcomes?
  - Evaluate model behavior to assist verification/validation
  - To assist in process of model development
- 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.)

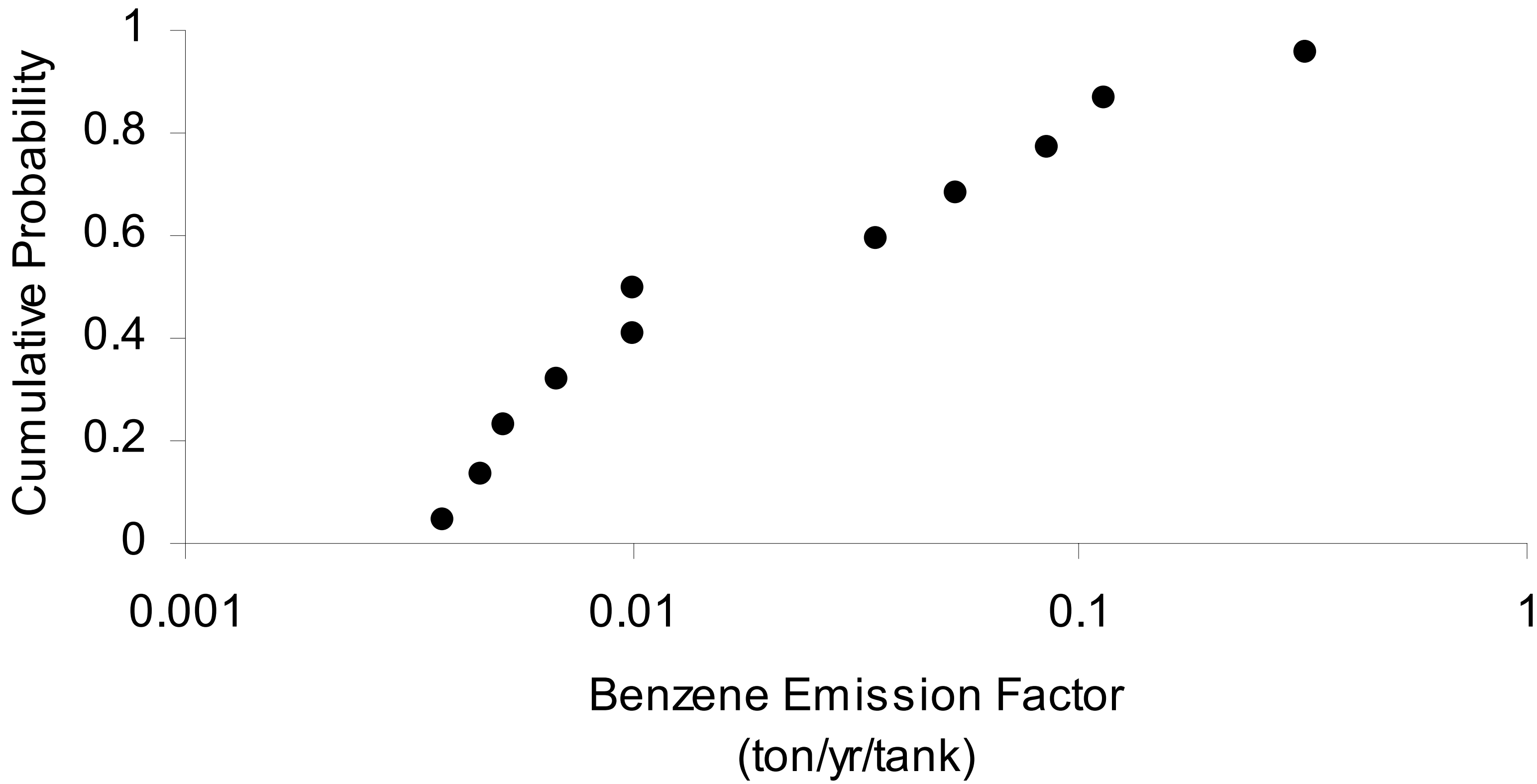
Local Sensitivity Analysis



Summary of Probabilistic Emissions Case Studies at NCSU

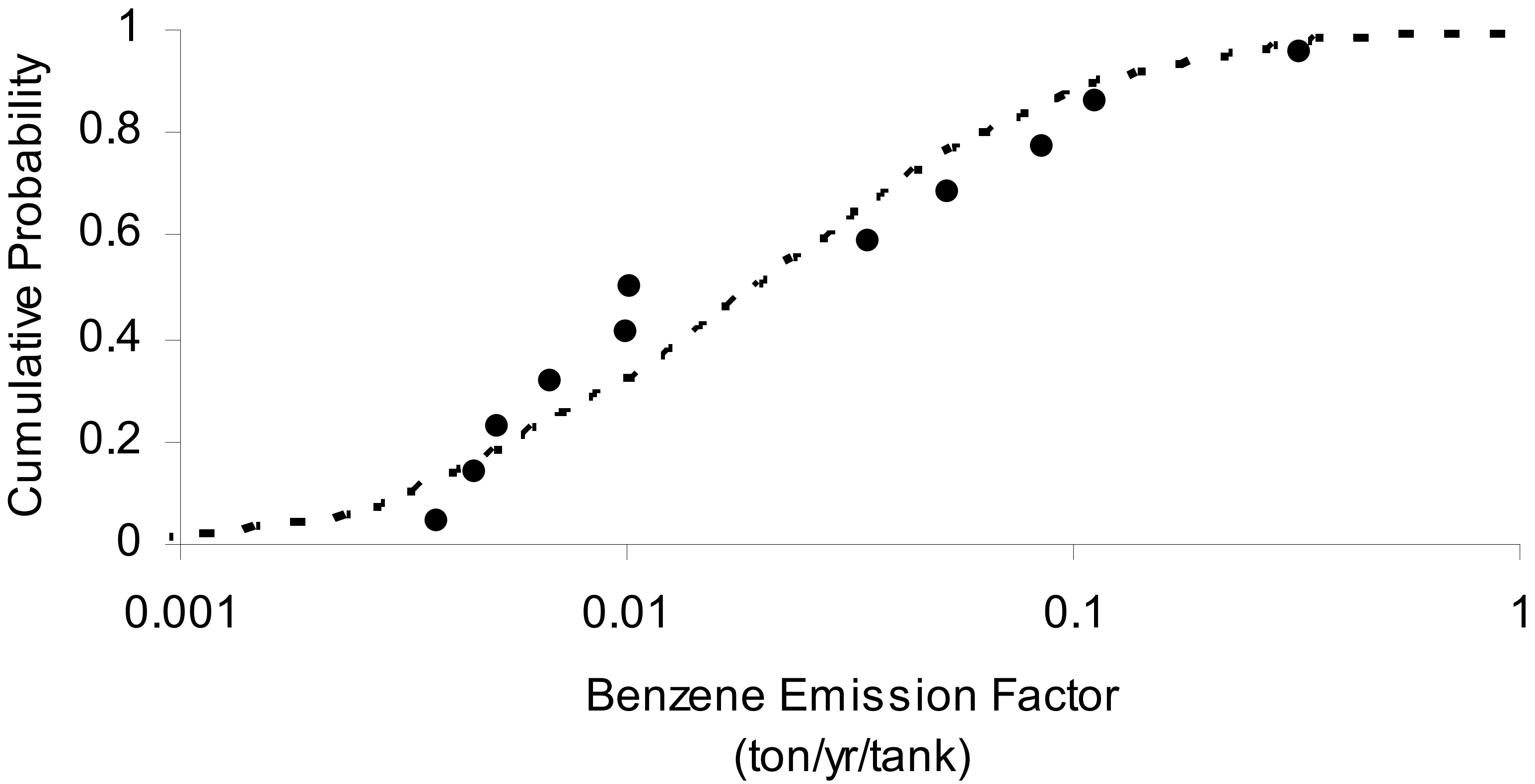
- Case Studies (examples):
  - Point sources
    - Power Plants
    - Natural gas-fired engines (e.g., compressor stations)
  - Mobile sources
    - On-Road Highway Vehicles
    - Non-Road Vehicles (e.g., Lawn & Garden, Construction, Farm, & Industrial)
  - Area sources
    - Consumer/Commercial Product Use
    - Natural Gas-Fueled Internal Combustion Engines
    - Gasoline Terminal Loading Loss
    - Cutback Asphalt Paving
    - Architectural Coatings
    - Wood Furniture Coatings
- Pollutants
  - NO_x
  - VOC
  - Urban air toxics (e.g., Houston case study)

Example of Benzene Emission Factor for Summer Storage Losses at a Storage Tank: Empirical Distribution

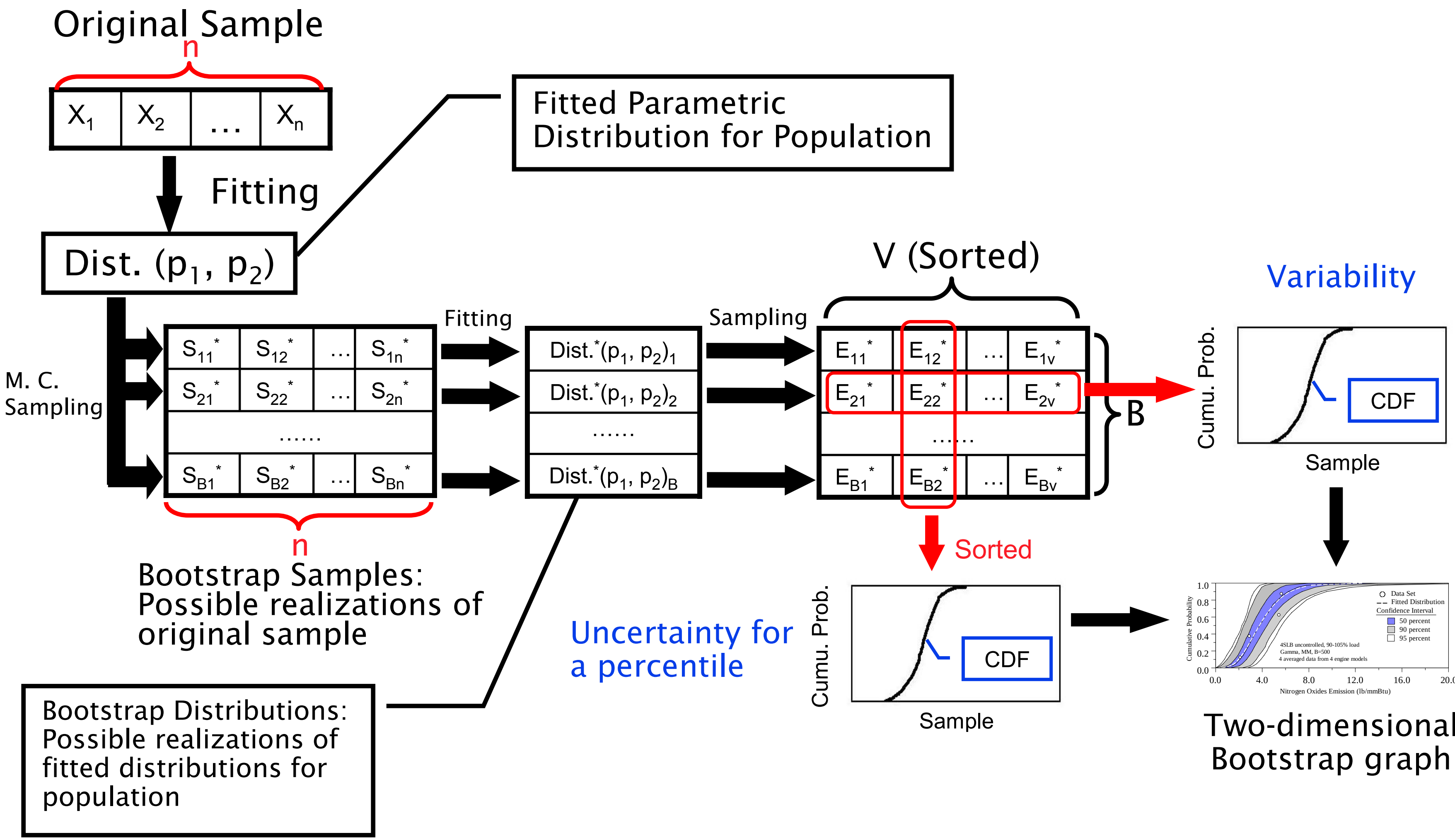


NC STATE UNIVERSITY

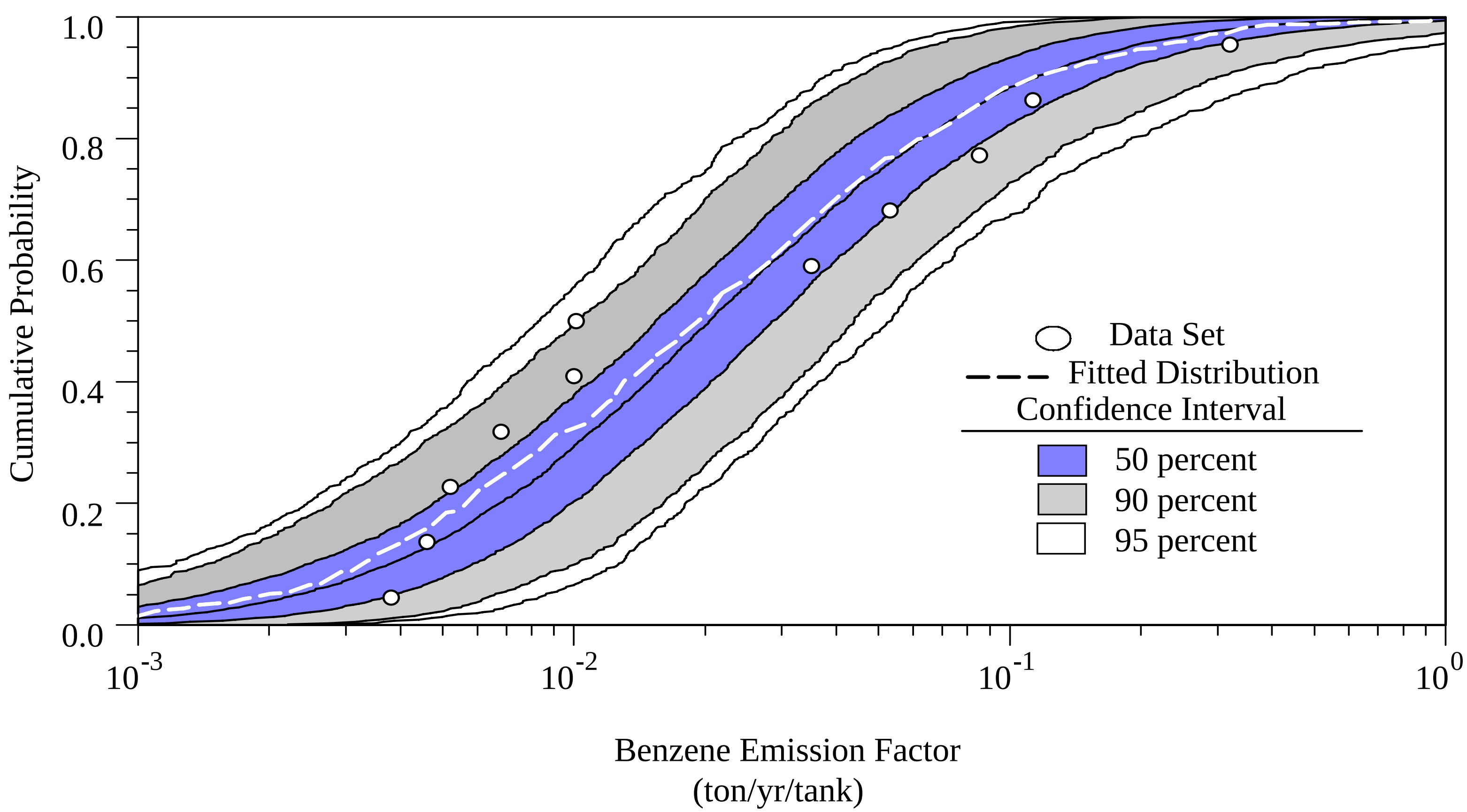
Example of Benzene Emission Factor: Fitted Lognormal Distribution



Bootstrap Simulation Process

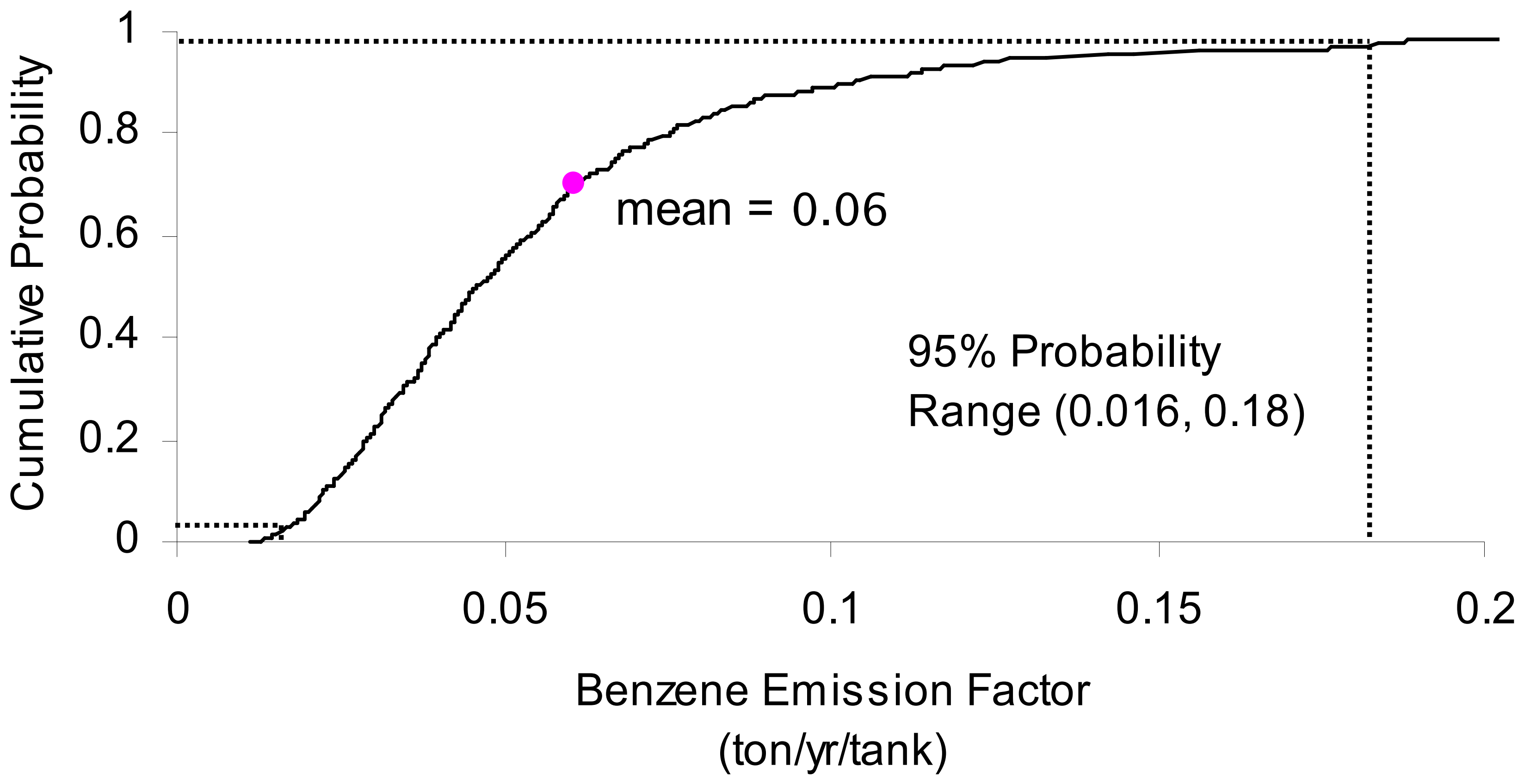


Example of Benzene Emission Factor: Confidence Interval for the Fitted Distribution



NC STATE UNIVERSITY

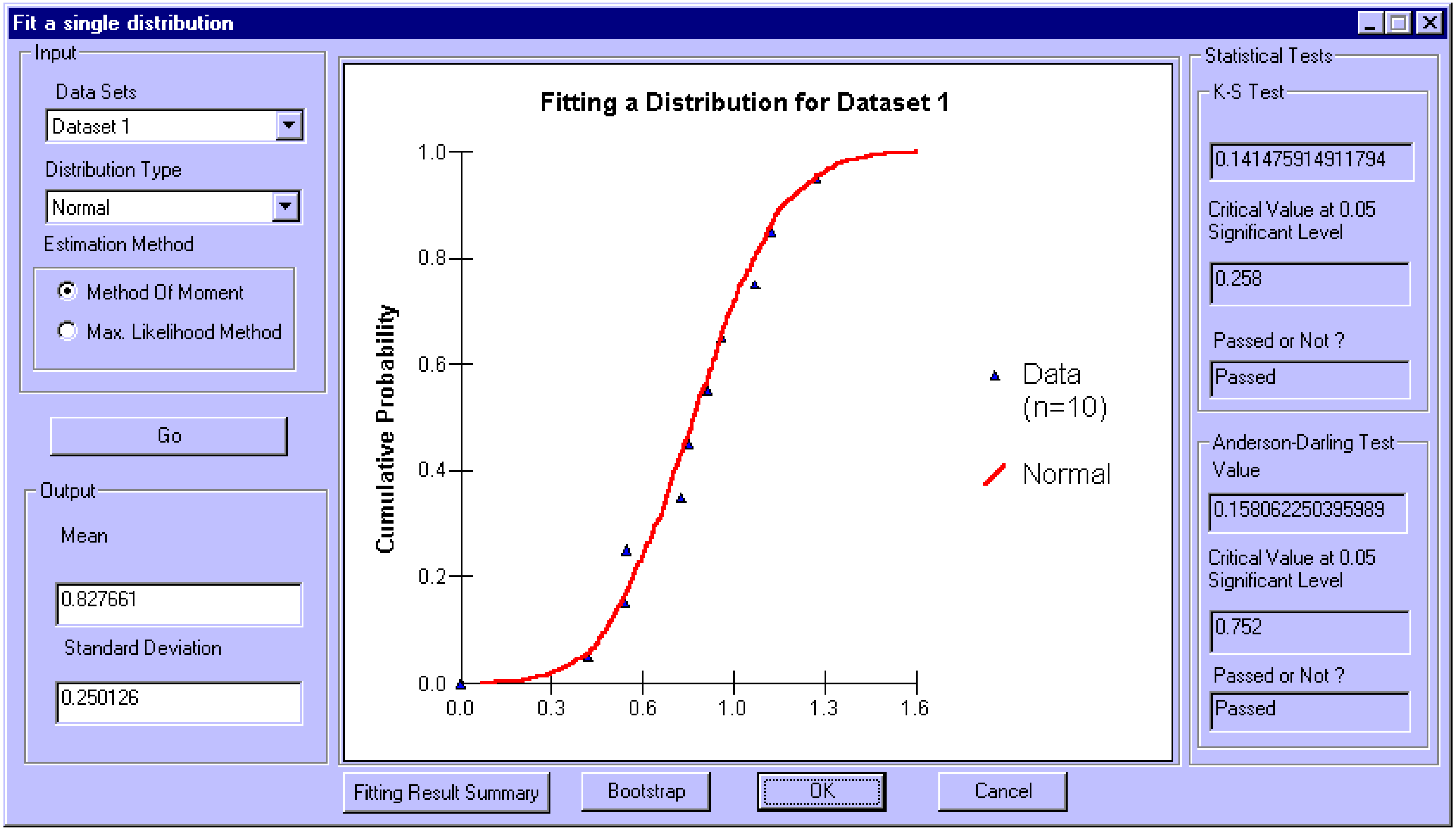
Example of Benzene Emission Factor: Uncertainty in the Mean



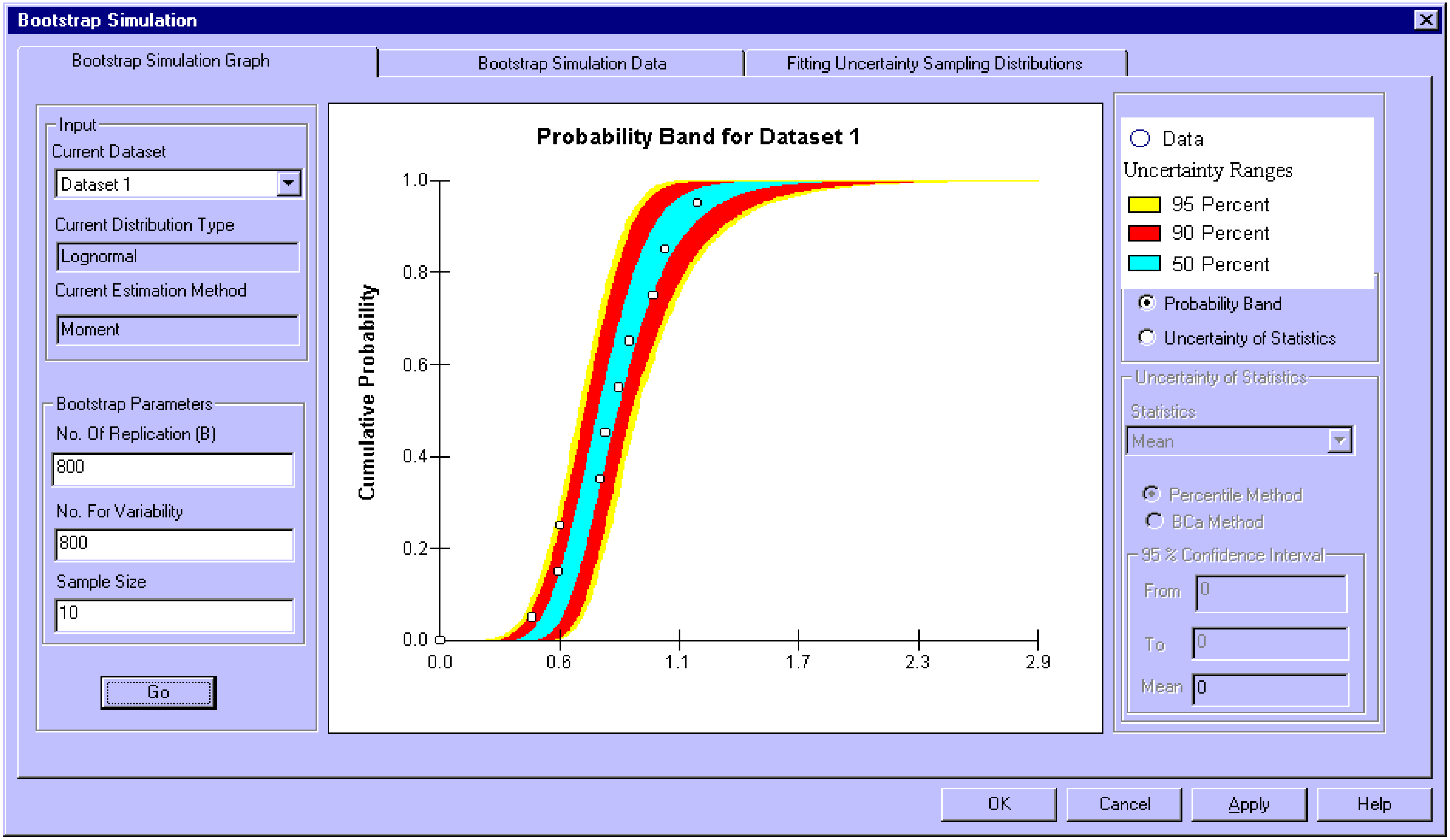
Uncertainty in mean -73% to +200%

NC STATE UNIVERSITY

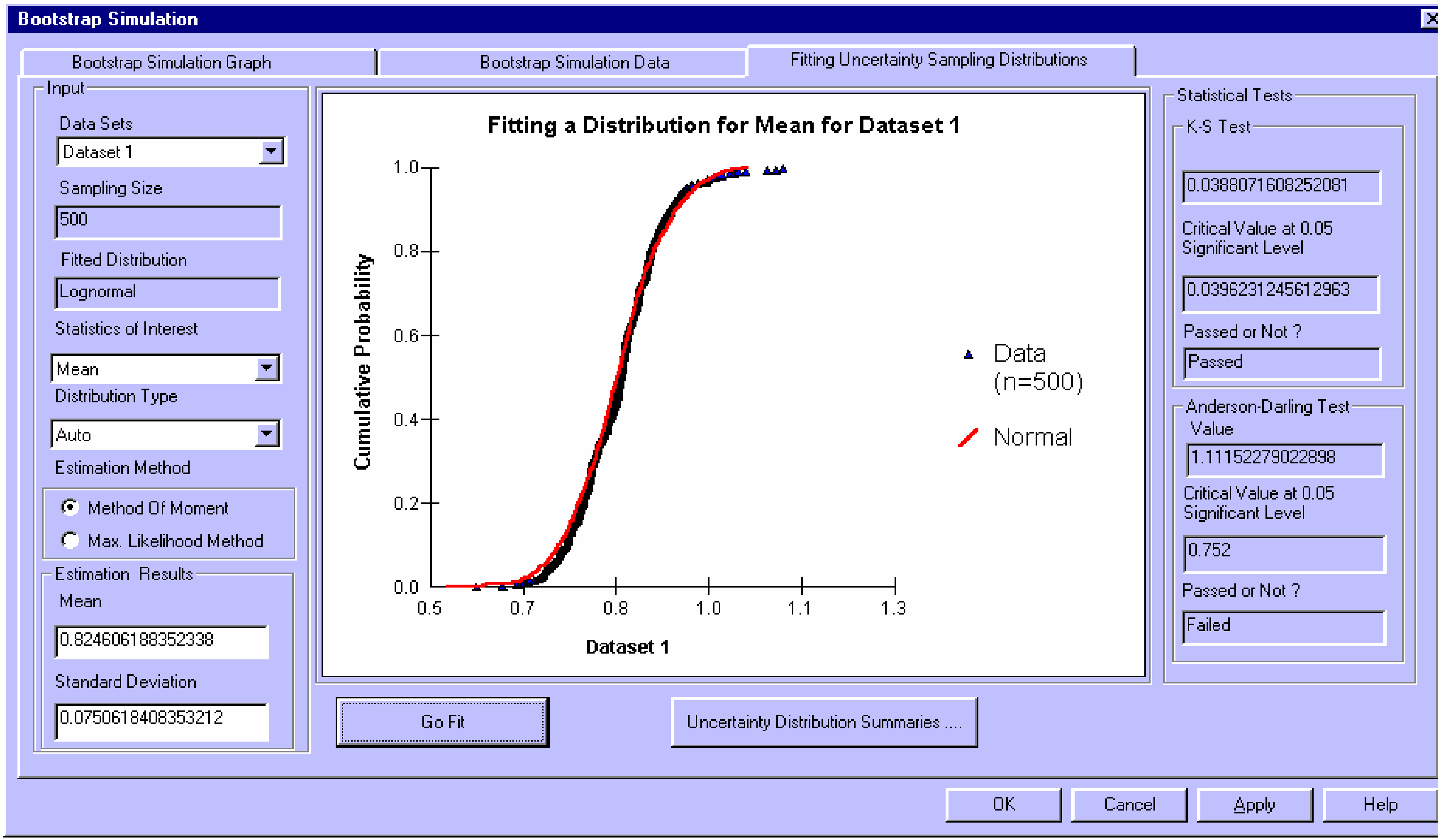
Using AuvTool to Fit a Distribution for Variability



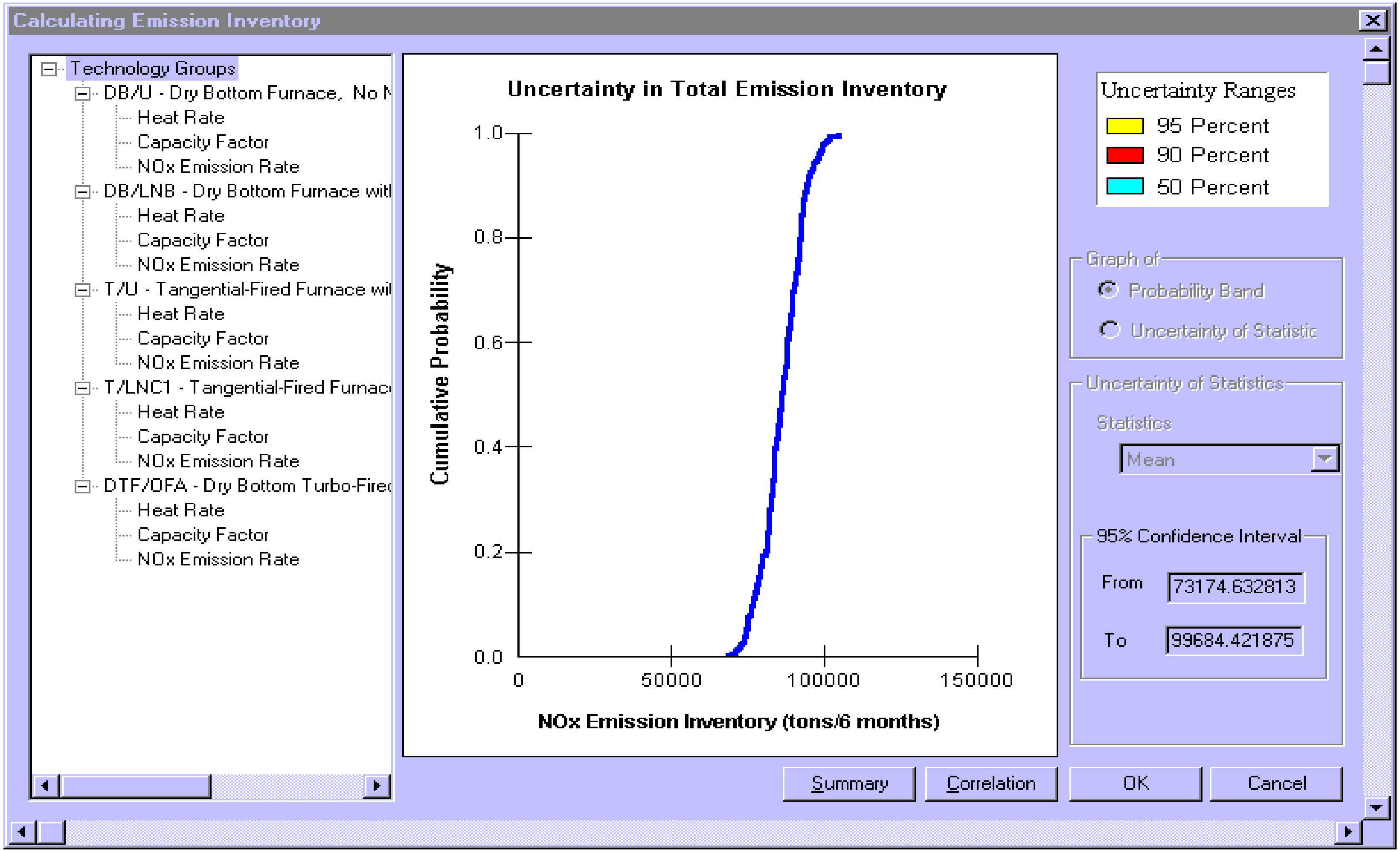
Using AuvTool for Bootstrap Simulation



Using AuvTool to Quantify Uncertainty in the Mean



Uncertainty in Total Emission Inventory: AUVEE Prototype Software



Summary of Probabilistic Emission Inventory

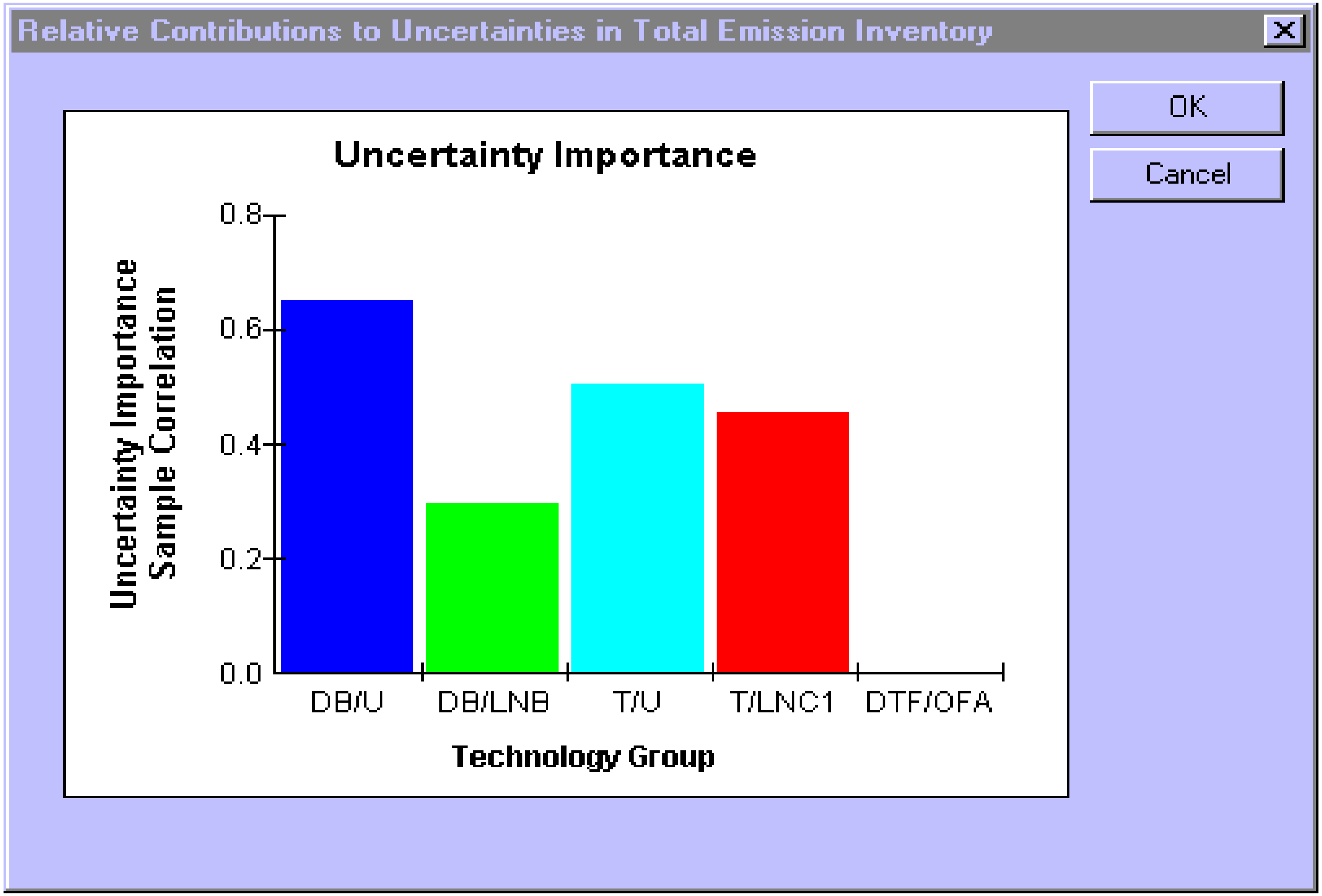
Summary Table of Emission Inventories

Technology Groups	Mean	2.5 Percentile	97.5 Percentile
DB/U	31960.00	22804.80	41827.08
DB/LNB	8198.76	5771.43	10815.03
T/U	20745.48	14944.52	28123.16
T/LNC1	25437.75	20332.91	31583.65
DTF/OFA	0.00	0.00	0.00
Total Inventory	86341.99	75067.33	99319.89

OK

NC STATE UNIVERSITY

Identification of Key Sources of Uncertainty in an Inventory



Summary of Probabilistic Emission Inventories for Selected Air Toxics

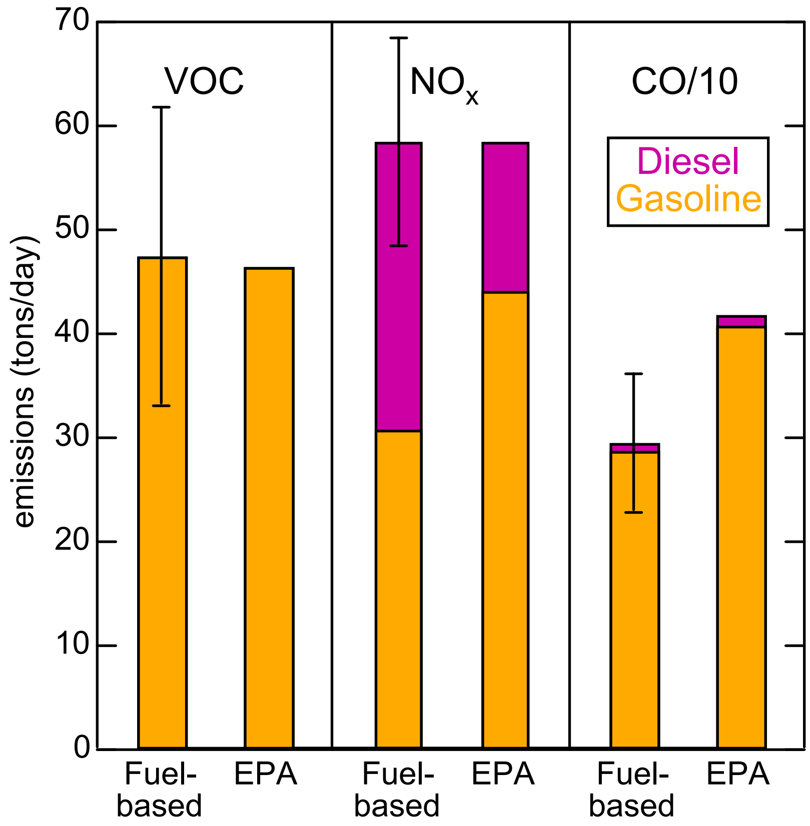
City	Pollutant	95% Probability Range Relative to the Mean Estimate (%, %)
Houston	Benzene	(-46, 108)
	Formaldehyde	(-35, 67)
	Chromium	(-20, 34)
	Arsenic	(-69, 203)
Jacksonville	1,3-butadiene	(-46, 108)
	Mercury	(-25, 30)
	Arsenic	(-83, 243)
	Benzene	(-56, 146)
	Formaldehyde	(-42, 89)
	Lead	(-54, 175)

NC STATE UNIVERSITY

Key Sources of Uncertainty

City	Pollutant	Key Sources of Uncertainty **(number of dominant sources)**
Houston	Benzene	Gasoline onroad mobile sources **(1)**
	Formaldehyde	Onroad mobile sources; Nonroad mobile sources **(2)**
	Chromium	Chemical manufacturing-fuel fired heaters **(1)**
		External utility coal combustion boilers; Hard chromium electroplating **(2)**
	Arsenic	External coal combustion utility boilers **(1)**
Jacksonville	1, 3-butadiene	Onroad mobile sources **(1)**
	Mercury	External coal combustion boilers **(1)**
	Arsenic	External coal combustion boilers **(1)**
	Benzene	Onroad mobile source **(1)**
	Formaldehyde	Onroad mobile source; Aircraft **(2)**
	Lead	External coal combustion boilers; Fuel oil external combustion; Waste oil external combustion **(3)**

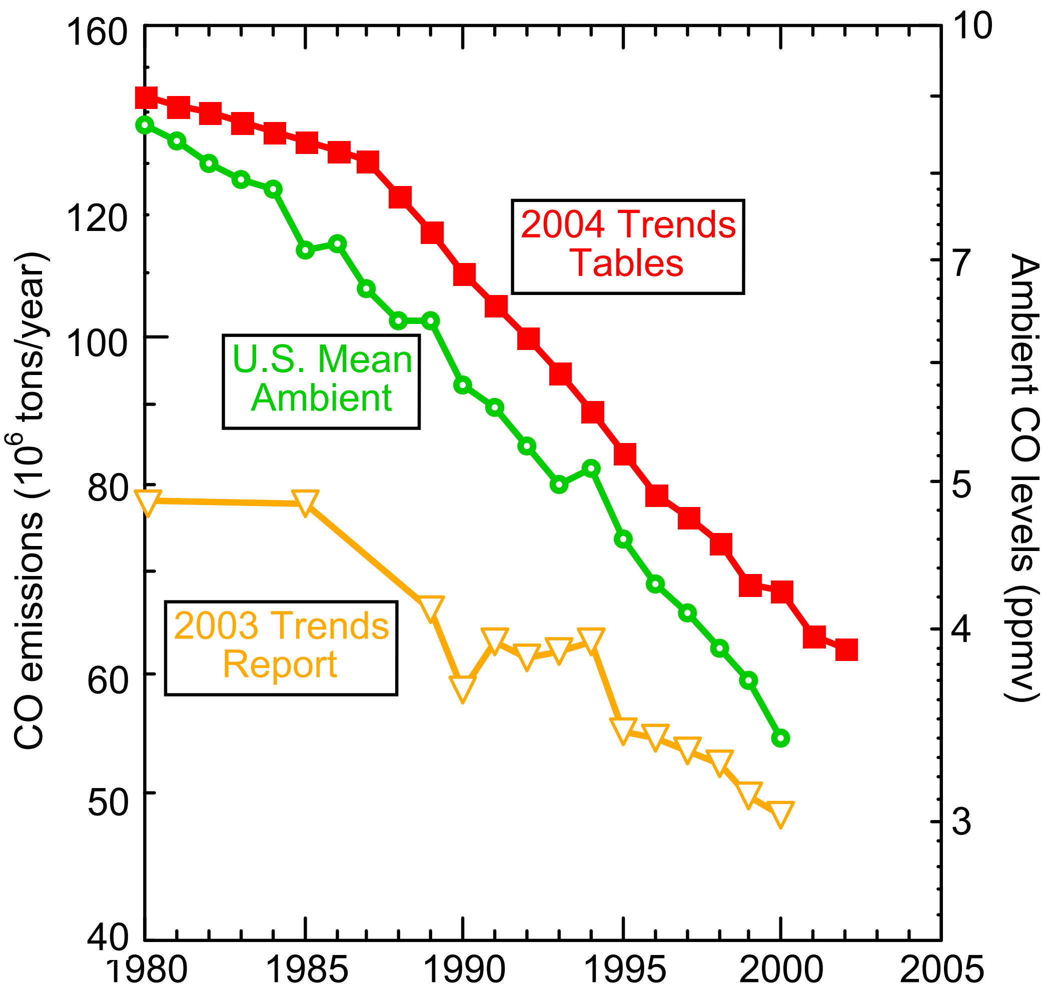
Comparison of Fuel-Based Versus Mileage-Based Mobile Source Inventories



Source: NARSTO (2005)

NC STATE UNIVERSITY

Example: Comparing Trends in Estimated Emissions and Ambient Monitoring Data



Source: NARSTO (2005)

Emission Factors and Inventories: Findings (1)

- Visualization of data informs choices of distribution models
- Difficult to find the original data (and its documentation) used by EPA and others.
- Test methods for (some) sources are not representative of real world operation - need for real world data, comparison of top-down vs. bottom-up approaches, and/or expert judgment to improve representativeness

Emission Factors and Inventories: Findings (2)

- Need systematic reporting of the precision and accuracy of measurement/test methods
- Uncertainties in emission factors are typically positively skewed, unless the uncertainties are relatively small (e.g., less than about plus or minus 30 percent)
- The quantifiable portion of uncertainty attributable to random sampling error can be large and should be accounted for when using emission factors and inventories

Emission Factors and Inventories: Findings (3)

- Range of variability and uncertainty is typically much greater as the averaging time decreases
- Even for sources with continuous emissions monitoring data, there is uncertainty regarding predictions of future emissions
- Prototype software demonstrates the feasibility of increasing the convenience of performing probabilistic analysis
- Uncertainties in total inventories are often attributable to just a few key emission sources

General Recommendations (1)

- Uncertainty and sensitivity analysis should be used to answer key decision maker and stakeholder questions, e.g.,:
  - Identify realistic emission estimates
  - Characterize variability in emission to help guide development of control strategies
  - prioritize scarce resources toward additional research or data collection
  - make choices among alternatives in the face of uncertainty,
  - evaluate trends over time, and so on.
- Uncertainty and sensitivity analysis should be incorporated into model and input data development
- There should be minimum reporting requirements for uncertainty in data (e.g., summary statistics such as mean, standard deviation, sample size)

General Recommendations (2)

- Flexibility - there are many possible approaches to analysis of uncertainty and sensitivity.
- Human resources should be appropriately committed.
  - Adequate time and budget to do the job right the first time (could save time and money in the long run)
  - Adequate training and peer review
  - Workshops and other training opportunities, and periodic refinement of authoritative compilations of techniques and recommended practice

NC STATE UNIVERSITY

General Recommendations (3)

- Software tools substantially facilitate both uncertainty and sensitivity analysis and should be further developed
- Further work - best techniques for communication, real-world information needs for decision makers
- Compilation of relevant case studies and insight

For More Information

- http://www4.ncsu.edu/~frey/
- Go to *Publications* link
- Many conference papers and technical reports on this subject can be downloaded
- Numerous journal papers on this topic are listed with links to where to obtain these papers (or ask me for a reprint)
- frey@ncsu.edu