Bootstrap Simulation and Two-Dimensional Monte Carlo Simulation: Dealing with Variability and Uncertainty, Mixture Distributions, Measurement Error, and Censored Data

H. Christopher Frey
Junyu Zheng

Department of Civil Engineering
North Carolina State University

Society for Risk Analysis Annual Meeting
Sunday, December 8, 2001
New Orleans, LA

Workshop Description

This half-day workshop presents the methodological issues for dealing with variability and uncertainty using bootstrap simulation and two-dimensional Monte Carlo Simulation. These issues include: What do you do when you have a data set that represents variability in a quantity? Should you fit a parametric distribution (e.g., normal, lognormal, gamma, Weibull, beta) to the data? If so, how should you choose such a distribution? Should you use the data to describe an empirical distribution? How can you quantify uncertainty in a statistic of the data or the distribution, such as the mean and standard deviation? How can you quantify uncertainty in the parameters of a distribution fitted to the data? What can you do when your data are actually based upon a mixture of distributions that cannot be well-described by a single component distribution? What if you know that the observed data contain a substantial amount of measurement error? What if some of your data are below a detection limit? This workshop will answer these questions. The methods, case studies, and software that will be presented in this workshop are based upon several years of research at NC State regarding development and demonstration of quantitative methods for characterizing variability and uncertainty in the inputs to emission inventories and risk assessment models. One of the outcomes of this research is a Windows98-based software tool, with a graphical user interface, for fitting distributions to data and characterizing uncertainty in statistics of the fitted single component distributions using bootstrap simulation. The workshop will feature an introduction to the theory and practice of fitting distributions to data, using bootstrap simulation to quantify uncertainty in the fitted distributions, and how to handle specific situations involving mixture distributions, measurement errors and censored data. Workshop participants will be provided with course notes and with a copy of the AuvTool software, including electronic versions of the technical and user's manual. The methods will be illustrated with case study examples. This workshop is aimed at practitioners who are already familiar with basic concepts of Monte Carlo simulation and who wish to refine their knowledge regarding methods for dealing with both variability and uncertainty in the context of a two-dimensional simulation.

Background of Presenters

H. Christopher Frey, Ph.D.
Associate Professor of Civil Engineering
North Carolina State University
Campus Box 7908
Raleigh, NC 27695-7908
Tel: 919-515-1155
Email: frey@eos.ncsu.edu

Dr. H. Christopher Frey is an Associate Professor of Civil Engineering at North Carolina State University. Since joining the faculty at NC State in January of 1994, Dr. Frey has been the principal investigator for 14 sponsored research projects and has been a co-principal investigator on three others. His projects are primarily in the areas of: (1) measurement and modeling of highway vehicle emissions; (2) modeling and assessment of advanced air pollution control and prevention systems; and (3) quantification of variability and uncertainty in air pollutant emission factors and emission inventories. Sponsors of Dr. Frey's research program at NC State have included NC Department of Transportation, Center for Transportation and the Environment, Transportation Research Board, MCNC, EPRI, U.S. Department of Energy, U.S. Environmental Protection Agency, and the National Science Foundation.

Dr. Frey's projects in the area of measurement and modeling of highway vehicle emissions have included five projects involving development of probabilistic emission factor estimates for light duty gasoline vehicles and for off-road vehicles. Dr. Frey has led two projects involving remote sensing of on-road tailpipe emissions of carbon monoxide (CO) and hydrocarbons (HC) for school buses, transit buses and light duty vehicles. The most recent of the two remote sensing projects was sponsored by the TRB ITS-IDEA program and was done in collaboration with Dr. Nagui Rouphail and the National Institute of Statistical Sciences. The collaborative project involved simultaneous measurement of on-road vehicle emissions using remote sensing and traffic data using a video-based area wide traffic detector. The simultaneously collected traffic and emissions data were used to develop statistical insights into vehicle emissions based upon observable traffic variables. As part of a recent project for the Office of Air Quality Planning and Standards for the U.S. Environmental Protection Agency, Dr. Frey has revisited and expanded upon the earlier work regarding probabilistic modeling of highway vehicle emission factors. The most recent project, completed in January 2002 for EPA's Office of Transportation and Air Quality (OTAQ), involved demonstration of a conceptual approach for developing vehicle emission factors from on-board data using a modal approach. A common theme of all of Dr. Frey's highway vehicle emissions projects is the application of statistical techniques to quantify both variability and uncertainty in emissions.

Dr. Frey is currently the Principal Investigator for two recently initiated projects involving quantification of variability and uncertainty in emission inventories of hazardous air pollutants, nitrogen oxides, and volatile organic compounds. These two projects, sponsored by the National Center for Environmental Research and Quality Assurance (NCERQA) of the U.S. EPA, include development of emission inventories for highway vehicles.

Dr. Frey was co-principal investigator on NCDOT sponsored Research Project 99-8 to evaluate the impact of traffic signal timing and coordination on real-world vehicle emissions. This project involved the first application of the OEM-2100 instrument for measuring vehicle tailpipe emissions with an on-board system. The project has successfully demonstrated the use of the OEM-2100 for collecting real world data and has resulted in the development of data collection protocols and study design methods that will be used in the proposed work.

Dr. Frey's contributions in environmental science and engineering have been recognized by national awards, such as a 1992 AAAS/EPA Environmental Science and Engineering Fellowship, a Faculty Early Career Development (CAREER) Grant by the National Science Foundation in 1997, and the 1999 Chauncey Starr award from the Society for Risk Analysis. Dr. Frey has been invited to participate on several national panels, including a U.S. EPA Science Advisory Board subcommittee on Residual Risk and several U.S. EPA workshops pertaining to the development and application of probabilistic methods in environmental science and engineering. In April 1998, Dr. Frey chaired a U.S. EPA workshop held in New York City on methods for the development and use of probability distributions to represent environmental data. On numerous occasions, Dr. Frey has been asked to provide peer review on major EPA documents, including technical reports, regulatory support documents, and Reports to Congress.

Dr. Frey has written approximately 120 journal papers, conference papers, and technical reports. In 1999 his first book, co-authored with Alison Cullen of the University of Washington, was published by Plenum Press on the topic of "Probabilistic Techniques in Exposure Assessment."

Dr. Frey earned a B.S. in Mechanical Engineering from the University of Virginia, and both a M.E. in Mechanical Engineering and a Ph.D. in Engineering and Public Policy from Carnegie Mellon University. His experience prior to joining the faculty at NC State includes a position as an Environmental Engineer at Radian Corporation from 1987 to 1988 and 2.5 years on the research faculty at Carnegie Mellon after completing his doctorate.

Junyu Zheng, Ph.D.
Post-doc Research Fellow of Civil Engineering
North Carolina State University
Campus Box 7908
Raleigh, NC 27695-7908
Tel: 919-515-4232
Email: jzheng3@eos.ncsu.edu

Dr. Junyu Zheng is now a post-doctoral research fellow in the Department of Civil Engineering at NCSU, and he conducts his research under the supervision of Dr. Frey in the Computational Laboratory for Energy, Air and Risk (CLEAR), which is part of the Environmental Engineering program within the department. His Ph.D. research focused on the application of statistics and computer techniques to the quantification of variability and uncertainty in emission estimation, exposure assessment and probabilistic modeling. In collaboration with Dr. Frey, Dr. Zheng has completed two journal papers regarding quantification of variability and uncertainty in power plant NO_x emissions and highway vehicle emissions, respectively. Dr. Zheng and Dr. Frey are

working together on two additional papers regarding methods for quantifying variability and uncertainty for mixture distributions and methods for accounting for measurement error in individual data values when fitting distributions to data. These four papers comprised Dr. Zheng's Ph.D. dissertation.

Dr. Junyu Zheng has done extensive work regarding quantification of variability and uncertainty in highway emission factors based upon Mobile5b and has developed software tools for statistical analysis of environmental data featuring the use of bootstrap simulation. In the study of uncertainty in Mobile5b emission factors, Dr. Zheng, in collaboration with Dr. Frey, has developed a methodology for quantifying inter-vehicle variability and fleet average uncertainty in the temperature and Reid vapor pressure factors. Numerical methods based upon the use of Monte Carlo simulation, Latin Hypercube and bootstrap simulation were employed. The case studies of CO, HC and NO_x emission for light duty gasoline vehicles were used to demonstrate the use of the approach.

During his Ph.D. study, Dr. Zheng developed two software tools AuvTool (Analysis of Uncertainty and Variability Tool) and AUVEE (Analysis of Uncertainty and Variability in Emission Estimation). The two tools implement the relevant methodologies for dealing with variability and uncertainty, featuring the use of bootstrap simulation and two-dimensional Monte Carlo Simulation. AuvTool can be generally used in any quantitative analysis field for characterizing variability and uncertainty in model inputs.

Dr. Zheng is now working on uncertainty analysis of the EPA SHEDS (Stochastic Human Exposure Dose Simualtion)/Pesticides model. Part of his work is to incorporate the AuvTool into the EPA/SHEDS models.

Dr. Junyu Zheng earned a Master of Science in Environmental Engineering from Tsinghua University of China in 1996. Tsinghua University is highly respected and is considered to be the top engineering university in the China. He received his Bachelor's degree from Wuhan Urban Construction Institute in 1991.

Preliminary Workshop Agenda

1:00 – 1:05 Welcome and Introduction Materials

1:05 – 2:30 Quantification of Variability and Uncertainty
- Statistical estimation
- Empirical distribution
- Parametric distributions
- Sampling distribution and confidence intervals
- Analytical methods – advantages, disadvantages, and limitations.
- Numerical methods – bootstrap simulation
- Parameter estimation
- Two-Dimensional Monte Carlo Simulation
- Example case studies
- Goodness-of-fit tests

2:30 – 3:00 Introduction to AuvTool, Installation and its Use
- AuvTool Design
- AuvTool Main Features
- AuvTool Installation
- Demonstration of AuvTool

3:00 – 3:15 break

3:15 – 3: 45 Censored Data
- Brief overview of theory for fitting distributions to censored data
  - One detection limit
  - Multiple detection limits
- Use of bootstrap pairs to quantify uncertainty in the statistics of distributions fit to censored data
- Properties of general solutions for variability and uncertainty applied to censored data
- Example case studies

3:45 – 4:15 Mixture Distributions
- When are mixture distributions useful or needed?
- Parameter estimation of mixture distributions
- Quantification of uncertainty in the parameters of mixture distributions
- Properties of mixture distributions
- Example Case Studies

4:15 – 4:45 Measurement Error
- Brief overview of methods for dealing with measurement error
  - Deconvolution methods
  - Simpler techniques

- Explanation of a practical technique for dealing with additive measurement errors
- Use of bootstrap pairs to address variability and uncertainty for data sets that include measurement error
- General properties of solutions for variability and uncertainty
- Example case studies

4:45 – 5:00 Summary, Final Discussion and Evaluations