ABSTRACT

Zhao, Yuchao. Probabilistic Modeling of Variability and Uncertainty in Urban Air Toxics Emissions (Under the direction of Dr. H. Christopher Frey)

Air toxic emission factor data often contain one or more censored points below a single or multiple detection limits. Such data sets are referred to as "censored." Conventional methods used to deal with censored data sets include removing non-detects, or replacing the censored points with zero, half of the detection limit or the detection limit. However, the estimated means of the censored data set by conventional methods are usually biased. Here, an approach to quantification of the variability and uncertainty of censored data sets is demonstrated. Empirical bootstrap simulation is used to simulate censored bootstrap samples from the original data. Maximum Likelihood Estimation (MLE) is used to fit parametric probability distributions to each bootstrap sample, thereby specifying alternative estimates of the unknown population distribution of the censored data sets. Sampling distributions for uncertainty in statistics such as the mean, median and percentile are calculated. The robustness of the method was tested by application to different degrees of censoring, sample sizes, coefficients of variation and numbers of detection limits. Lognormal, gamma and Weibull distributions were evaluated. The reliability of using this method to estimate the mean is proved. The application of MLE/Bootstrap was compared favorably to results obtained with the non-parametric Kaplan-Meier method, which verify the accuracy of this method.

The MLE/bootstrap method is applied to 16 cases of censored air toxic emission factors, including benzene, formaldehyde, Benzo(a)pyrene, mercury, arsenic, cadmium, total chromium, chromium VI and lead with single or multiple detection limits from coal, fuel oil and/or wood waste external combustion sources. The data differs regarding sample size, censoring degree,

inter-unit variability and so on. The proportion of censored values in the emission factor data ranges from 4 to 80 percent. The largest range of uncertainty in the mean was obtained for the external coal combustion benzene emission factor, with a 95 percent probability range of minus 93 to plus 411 percent of the mean.

Probabilistic emission inventories were developed for benzene, formaldehyde, chromium, and arsenic for Houston 1996 emission inventory and for 1, 3-butadiene, mercury, arsenic, benzene, formaldehyde and lead. Parametric distributions for inter-unit variability were fit using maximum likelihood estimation (MLE) and uncertainty in mean emission factors was estimated using parametric bootstrap simulation. For data sets containing one or more non-detected values, empirical bootstrap simulation was used to randomly sample detection limits for non-detected values and observations for sample values, and parametric distribution for variability were fit using MLE estimators for censored data. Goodness-of-fit for censored data was evaluated using the Kolmogorov-Smirnov test applied to a modified data set and by comparison of cumulative distributions of bootstrap confidence intervals and empirical data. The emission inventory 95 percent uncertainty ranges are as small as minus 25 to plus 42 percent for chromium for Houston to minus 75 to plus 224 percent for arsenic for Jacksonville. Uncertainty was dominated by only a few source categories. Recommendations are made for future improvements to the analysis.

PROBABILISTIC MODELING OF VARIABILITY AND UNCERTAINTY IN URBAN AIR TOXICS EMISSIONS

by

YUCHAO ZHAO

A dissertation submitted to the Graduate Faculty of
North Carolina State University
In partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Civil Engineering

Raleigh

2003

Approved BY:

Chair of Advisory Committee

To my late grandmother and my family

BIOGRAPHY

Yuchao Zhao received her bachelor degree of environmental engineering in the Department of Environmental Science and Engineering from Tsinghua University, Beijing, China, in 1997. The topic of her B.S. thesis is Study of Microbial Coal Desulphurization. Yuchao Zhao gained the honor of entering the graduate program in the department directly, waived of the admission test in 1997. She worked as research assistant at the Environmental Planning and Management Lab and her research focused on uncertainty analysis of plume models. She earned her Master degree of environmental engineering in the Department of Environmental Science and Engineering from Tsinghua University in 2000. Yuchao Zhao is very grateful to the eight year life in Tsinghua University.

Yuchao Zhao came to North Carolina State University, Raleigh, North Carolina, USA, in August 2000 to pursue her Ph.D. degree in Environmental Engineering program. During her Ph.D study, she worked as a research assistant at the Computational Laboratory for Energy, Air and Risk (CLEAR).

Yuchao Zhao's research interests include application of mathematics/statistics method and computer techniques to develop model, probabilistic risk analysis, uncertainty analysis and sensitivity study.

ACKNOWLEDGEMENTS

This research was supported by U.S. EPA STAR Grants Nos. R826790.

Part of the work was funded via the Carolina Environmental Program.

The author appreciates Dr. H. Christopher Frey for his constant inspiration and guidance throughout the course of this research. Appreciation is also extended to Drs. M. –S. Yim, E.D. Brill and D. Vandervaart for their valuable suggestions. The author appreciates the suggestions of Dr. Daowen Zhang on this work.

The author thanks all other members in the group of Computational Laboratory for Energy, Air and Risk.

TABLE OF CONTENTS

LIST OF TABLES……viii
LIST OF FIGURES…..……..……..……..……...……...….x

INTRODUCTION..... 1

1.0 Introduction..... 2

1.1 Urban Air Toxics Emissions..……...2
1.2 Variability and Uncertainty…….……..3
1.3 Variability and Uncertainty in Censored Data……..4
1.4 Probabilistic Urban Air Toxics Emission Inventories……..9
1.5 Objectives……10
1.6 Overview of Research……..11
1.7 Organization……12
1.8 References……...13

QUANTIFICATION OF VARIABILITY AND UNCERTIANTY FOR CENSORED DATA SETS AND APPLICATION TO AIR TOXIC EMISSION FACTORS……..17

1.0 Introduction.……18

2.0 MLE Parameter Estimates for Censored Lognormal, Gamma and Weibull Distributions…….20

3.0 Estimation Uncertainty in Statistics Using Bootstrap Simulation..... 23

4.0 Robustness Test of the Bootstrap Method for Quantifying Uncertainty in the Statistics of Censored Data Sets 25

5.0 Evaluation of the Bootstrap Method for Estimating the Mean of Censored Data Sets 31

6.0 Verification of the Uncertainty Estimates from the Bootstrap Method with the Kaplan-Meier Estimator Method 32

7.0 Case Study for Estimation of Variability and Uncertainty in an Emission Factor Based Upon Censored Data 34

8.0 Conclusions 36

9.0 Acknowledgements 37

10.0References…….37

UNCERTAINTY FOR DATA WITH NON-DETECTS: AIR TOXIC EMISSIONS FROM COMBUSTION 56

1.0 Introduction……...57
1.1 Sources of Variability and Uncertainty…….59
1.2 Variability and Uncertainty in Urban Air Toxic Emissions……...60

2.0 Methodology 60
2.1 Conventional Approaches to Dealing with Censored Data……61
2.2 Method of Maximum Likelihood Estimation……61
2.3 Lognormal, Gamma and Weibull Distributions……62
2.4 Bootstrap Simulation……..63
2.5 Evaluation of Goodness-of-fit……64

3.0 Case Studies 65
3.1 Data……65
3.2 Quantification of Inter-Unit Variability and Graphical Evaluation of Goodness-of-fit……..66
3.3 Estimation of the Mean and Quantification of Its Uncertainty……..68
3.4 Comparison of Conventional and MLE/Bootstrap Methods for Mean Estimates……70
3.5 Comparison of Results for Censored Versus Modified Data……71

4.0 Conclusions and Recommendations 73

5.0 Acknowledgements 75

6.0 References 75

QUANTIFICATION OF VARIABILITY AND UNCERTINTY FOR AIR TOXIC EMISSION INVENTORIES WITH CENSORED EMISSIONS FACTOR DATA..... 87

1.0 Introduction..... 88

2.0 Methodology 90
2.1 Lognormal, Gamma and Weibull Distributions……..91
2.2 Goodness-of-fit Test……92
2.3 Monte Carlo Simulation of Uncertainty in the Emission Inventory Model……...93
2.4 Identification of Key Sources of Uncertainty……..93

3.0 Houston Emission Inventory..... 94

4.0 Results 96
4.1 Quantification of Variability and Uncertainty in Emission Factors……96
4.2 Development of Probabilistic Emission Inventories……98
4.3 Sensitivity Study to Identify the Key Sources of Uncertainty……99

5.0 Discussion..... 100

6.0 Acknowledgements 102

7.0 Supporting Information..... 103

8.0 References..... 103

DEVELOPMENT OF PROBABILISTIC EMISSION INVENTOTY OF AIR TOXICS FOR JACKSONVILLE, FLORIDA..... 145

1.0 Introduction..... 146

2.0 Methodology..... 148
2.1 Candidate Parametric Distributions for Inter-Unit Variability in Emission Factors……149
2.2 Maximum Likelihood Estimation and Method of Matching Moment……...149
2.3 Bootstrap Simulation……..150
2.4 Goodness-of-fit……...151
2.5 Monte Carlo Simulation of Uncertainty in the Emission Inventory Model...151
2.6 Sensitivity Analysis to Identify the Key Sources of Uncertainty……...152

3.0 Jacksonville Emission Inventory and Emission Factor Data..... 152

4.0 Results..... 154
4.1 Quantification of Variability and Uncertainty in Emission Factors……..154
4.2 Development of Probabilistic Emission Inventories…….158
4.3 Sensitivity Study to Identify Key Sources of Uncertainty……159

5.0 Discussion..... 160

6.0 Acknowledgements..... 163

7.0 References..... 163

CONCLUSIONS AND RECOMMENDATIONS..... 185

1.0 Conclusions and Recommendations..... 186
1.1 Conclusions Regarding MLE/Bootstrap Method for Censored Data……186
1.2 Conclusions Regarding the Development of Probabilistic Urban Air Toxics Emission Inventories for Houston, TX and Jacksonville, FL…..188
1.3 Recommendations for Future Work……...190

LIST OF TABLES

Part II

Table 1. Parameters of the Assumed Lognormal, Gamma and Weibull Population Distributions with Coefficients of Variation of 0.5, 1 and 2 41

Table 2. Detection Limits Associated with Different Cumulative Probabilities for Specified Population Distributions as a Function of Distribution Type and Coefficient of Variation 41

Table3. Average and 95 Percent Confidence Intervals of Mean Estimated from Bootstrap Simulation for Lognormal, Gamma and Weibull Distribution, Sample Sizes of 20, 40 and 100, Coefficients of Variation of 0.5, 1 and 2, and Censoring of 0, 30and 60 Percent……..42

Table 4. Average and 95 Percent Confidence Intervals of Mean Estimated from Bootstrap Simulation for Lognormal, Gamma and Weibull Distributions, Sample Sizes of 20, 40 and 100, Coefficients of Variation of 0.5, 1 and 2, and Multiple Detection Limits 44

Table 5. Average and 95 Percent Confidence Intervals of the Median and 90th Percentile Estimated from Bootstrap Simulation for the Weibull Distribution, Sample Sizes of 20 Sample sizes of 20 and 100, Coefficient of Variation = 2, and 0%, 30% and 60% Censoring 45

Table 6. Test of Hypothesis that the Average of Bootstrap Means Are Not Statistically Significantly Different from the Population Mean: Gamma Distribution, Population Mean = 1, Population Coefficient of Variation = 1, Sample Size = 20, Three Levels of Censoring 46

Table 7. Comparison of the Average Mean Value Based Upon 20 Synthetic Data Sets from a Gamma Distribution with Mean = 1, Coefficient of Variation = 1, and sample size = 20 for Three Levels of Censoring and Five Estimation Methodsa.... 47

Table 8. Input Data of Formaldehyde Emission Factors 47

PART III

Table 1. Summary of Emission Factor Data for Selected Urban Air Toxics from External Combustion Emission Source Categories 79

Table 2. Results from MLE/Bootstrap Simulation for Candidate Distributions Fit to Urban Air Toxic Emission Factor Data from External Combustion Sources..... 80

Table 3. Comparison of Estimated Means from Conventional and MLE/Bootstrap Methods.... 81

Table 4. Ratio of the Estimated Means and Width of the 95 Perrcent Confidence Intervals For the Means for the MLE/Bootstrap Method Applied to Modified Versus Censored Data. 82

PART IV

Table 1. Quantification of Variability and Uncertainty for Benzene Emission Inventory..... 109

Table 2. Quantification of Variability and Uncertainty for Formaldehyde Emission Inventory113

Table 3. Quantification of Variability and Uncertainty for Chromium Emission Inventory..... 115

Table 4. Quantification of Variability and Uncertainty for Arsenic Emission Inventory……..115

PART V

Table 1. Summary of Available Data for 1, 3-butadiene, Mercury, Arsenic, Benzene, Formaldehyde and Lead Emission Inventories 171

Table 2. Quantification of Variability and Uncertainty for 1, 3-butadiene Emission Inventory172

Table 3. Quantification of Variability and Uncertainty for Mercury Emission Inventory..... 174

Table 4. Quantification of Variability and Uncertainty for Arsenic Emission Inventory 175

Table 5. Quantification of Variability and Uncertainty for Benzene Emission Inventory..... 177

Table 6. Quantification of Variability and Uncertainty for Formaldehyde Emission Inventory.179

Table 7. Quantification of Variability and Uncertainty for Lead Emission Inventory..... 181

Table 8. Results of the Uncertainties in The Total Emission Inventories for 1, 3-butadiene, Mercury, Arsenic, Benzene, Formaldehyde and Lead..... 183

LIST OF FIGURES

Part II

Figure 1. Scheme of Quantification of Variability and Uncertainty for Censored Data Sets..... 48

Figure 2. 95 Percent Confidence Intervals for Uncertainty in the Cumulative Distribution Function for Lognormal Distribution with Coefficeint of Variation of 0.5 with Coefficient of Variation of 0.5, Sample Sizes of 10, 20 and 40, With No Censoring, 30% Censoring, and 60% Censoring (B=500) 49

Figure3. 95 Percent Confidence Intervals for Uncertainty in the Cumulative Distribution Function for Lognormal Distribution with Coefficient of Variation of 1, Sample Sizes of 10, 20 and 40, With No Censoring, 30% Censoring, and 60% Censoring (B=500)..... 50

Figure 4. 95 Percent Confidence Intervals for Uncertainty in the Cumulative Distribution Function for Lognormal Distribution with Coefficient of Variation of 2, Sample Sizes of 10, 20 and 40, With No Censoring, 30% Censoring, and 60% Censoring (B=500)..... 51

Figure 5. Comparison of 95 Percent Confidence Intervals of Cumulative Distribution Functions for the MLE/Bootstrap and Kaplan- Meier Estimator Methods for Different Amounts of Censoring and Coefficients of Variation for Lognormal Distributions……...52

Figure 6. Comparison of 95 Percent Confidence Intervals of Cumulative Distribution Functions for the MLE/Bootstrap and Kaplan-Meier Estimator Methods for Different Amounts of Censoring and Coefficients of Variation for Gamma Distributions 53

Figure 7. Comparison of 95 Percent Confidence Intervals of Cumulative Distribution Functions for the MLE/Bootstrap and Kaplan-Meier Estimator Methods for Different Amounts of Censoring and Coefficients of Variation for Weibull Distributions 54

Figure 8. Variability and Uncertainty in the Formaldehyde Emission Factor for a Combustion Source Estimated Based Upon a Lognormal..... 55

PART III

Figure 1. Variability and Uncertainty in Mercury Emission Factor from Coal Combustion Estimated Based Upon Three Distributions..... 84

Figure 2. Variability and Uncertainty in Cadmium Emission Factor from Coal Combustion Estimated Based Upon a Lognormal Distribution..... 85

Figure 3. Variability and Uncertainty in Arsenic Emission Factor from Fuel Oil Combustion Estimated Based Upon a Weibull Distribution 85

Figure 4. Variability and Uncertainty in Chromium Emission Factor from Coal Combustion Estimated Based Upon a Lognormal Distribution 86

Figure 5. Variability and Uncertainty in Formaldehyde Emission Factor from Fuel Oil Combustion Estimated Based Upon a Weibull Distribution 86

PART IV

Figure 1. Varaibility and Uncertainty in Chromium Emission Factor for Case 26 (Industrial boilers: residual oil) Estimated Based Upon a Lognormal Distribution……..120

Figure 2. Varaibility and Uncertainty in Chromium Emission Factor for Case 26 (Industrial boilers: residual oil) Estimated Based Upon a Gamma Distribution……...120

Figure 3. Varaibility and Uncertainty in Arsenic Emission Factor from Fuel Oil Combution Estimated Based Upon a Weibull Distribution.……..121

PART V

Figure 1. Variability and Uncertainty in Mercury Emission Factor from Pathological Waste Disposal Estimated Based Upon a Weibull Distribution 183

Figure 2. Uncertainty in 1, 3 - butadiene Emission Factor from Ships Estimated Based Upon Mixture Lognormal Distribution.……184

Figure 3. Variability and Uncertainty in Arsenic Emission Factor from Coal Combustion Estimated Based Upon a Lognormal Distribution 184

Part I

INTRODUCTION

1.0 Introduction

In this section, the motivations and backgrounds regarding development of probabilistic estimates of urban air toxics emissions are given. The primary objectives of this research are addressed. Finally, an overview of this research and the organization of this dissertation are introduced.

1.1 Urban Air Toxics Emissions

Toxic air pollutants, also known as hazardous air pollutants, are those pollutants that are known or suspected to cause cancer or other serious health effects, such as reproductive effects or birth defects, or adverse environmental effects. Examples of toxic air pollutants include benzene, which is often found is gasoline, and heavy metals, such as cadmium, mercury, chromium and lead.

The U.S. Environmental Protection Agency has developed an Integrated Urban Air Toxic Strategy, which includes a framework for addressing urban air toxics emissions. The 1990 Clean Air Act Amendments identified 189 HAPs and required maximum available control technology (MACT) on major sources of those chemicals. If the technology-based emissions controls are not sufficient, risk-based regulation is needed.[1] EPA developed a list of 33 urban air toxics, which represent the priority for additional risk assessment of the health effects of air toxics in urban areas.[1] In recent years, probabilistic risk assessment becomes more and more concerned. It provides quantitative information about the range and likelihood of risk analysis.

Urban air toxics emission inventories will be the basis for urban air toxic emission regulation. Emission inventories (EIs) are commonly obtained by the product of emission factors and activity factors. EIs are used by federal, state, and local governments and by private corporations for: (a) characterization of temporal emission trends; (b) emissions budgeting for

regulatory and compliance purpose; (c) prediction of ambient pollutant concentrations using air quality models;[2] and (d) development of exposure and risk analysis.[3]

1.2 Variability and Uncertainty

Variability is the heterogeneity of a quantity over time, space or members of a population.[3, 8] For example, a given human individual has a body weight, intake rate, lifetime, exposure duration, and activity pattern that are different from that of other individuals. Another example is that emissions of a particular pollutant typically differ from one specific emission source to another within a given source category. Variability can be represented by a frequency distribution showing the variation in a characteristic of interest over time, space.[2] Variability and uncertainty can both be represented as distributions in a two-dimensional modeling framework.[2, 4, 6] For example, confidence intervals can be constructed with respect to a best estimate of the cumulative distribution function (CDF) for variability. The range of the confidence intervals for the CDF represents uncertainty.

Uncertainty arises due to lack of knowledge regarding the true value of a quantity or regarding the true distribution for variability.[3, 8] Uncertainty in emissions is typically attributable to the following: (1) random measurement errors (lack of precision); (2) random sampling error; (3) systematic errors (bias or lack of "accuracy") such as would be caused by imprecise calibration or use of surrogate data (e.g., laboratory tests of vehicles rather than on-road measurements); and (4) human error such as random mistakes in entering or processing data. Uncertainty can be quantified as a probability distribution.[6]

Air toxic emissions are subject to both variability and uncertainty.[2, 4] If random errors and measurement error in the EIs are not quantified, erroneous inferences could be made regarding trends in emissions, source apportionment, compliance, the relationship between emissions and

ambient air quality and regulation decision based on risk analysis. Variability and uncertainty in air toxics emissions are a contributing factor to variability and uncertainty in estimates of exposure and risk. Quantification of variability and uncertainty in air toxics emissions is needed to identify high emitters or highly exposed populations as well as to characterize the quality of an emissions inventory and to target data collection to reduce uncertainty.

There is a growing track record of the motivations for characterizing variability and uncertainty applied to emission factors, emission inventories, air quality modeling, exposure assessment, and risk assessment. For example, the National Research Council (NRC) recommends that quantifiable uncertainties be addressed in estimating mobile source emission factors and logically this recommendation should be extended to other source categories.[7] The NRC has also addressed the need for quantification of uncertainties in emission inventories used in risk assessment.[8] Probabilistic techniques have been applied to estimate uncertainty in emission factors for mobile sources, major stationary sources and area sources, particularly for criteria pollutants (e.g., NO_x) and ozone precursors (e.g., volatile organic compounds).[9-16] Now, it is necessary and technically feasible to quantify variability and uncertainty in urban air toxic emissions.

1.3 Variability and Uncertainty in Censored Data

Because of inherent limitations of chemical/analytical measurement methods, urban air toxics emission data often contain several observations reported as below a detection limit (DL), and are usually referred to as "censored." [17] Many situations have a majority of measurements for urban air toxics that are below a detection limit.[18]

Generally, censored data include right-censored, left-censored and interval-censored data. For right-censoring, the data points larger than the detection limit are non-detects. For left-

censoring, the data points smaller than the detection limit are non-detects. For interval-censoring, the data points between two different detection limits are non-detects. For example, in many biomedical applications the primary endpoint of interest is time to a certain event, such as time to death. People are interested in characterizing the distribution of "time to event" for a given population. Typically, in biomedical applications the data are collected over a finite period of time and consequently the "time to event" may not be observed for all the individuals in the study sample.[19] This results in right-censored data. Environmental censored data are usually left-censored. For example, the non-parametric Kaplan-Meier estimator has been used to deal with left-censored pyrene concentration data collected from 20 Monitoring Stations in the Central Basin of Puget Sound from 1988-1990.[20] Furthermore, left-censored data are characteristic of many bioassays due to inherent limit of detection in the assays.[21] For example, plasma HIV RNA measurements collected for the Hemophilia Growth and Development Study are left censored data.[21] Studies on emergence times of teeth are also most often faced with interval-censored data.[22] The emergence time of a tooth is defined as the chronological age of a child when that tooth appears in the mouth.[22] When the precise emergence time is not available, it is given as interval censored data between two detection limits.

In statistics, the detection limit is a fixed value and a non-detect point x in a left-censored data set means $x \in [0, DL]$ and a non-detect point x in right-censored data set means $x \in [DL, +\infty)$. For example, in survival analysis, if the finite study time for event of "death" is 10 years, then the detection limit is 10 years. The event of "death" after 10 years is non-detected or censored. However, from an engineering and analytical chemistry perspective, the terminologies of "non-detect" and "detection limit" for measurements are more complicated than those in statistics.

Scientists from the International Organization for Standardization (ISO) and International Union of Pure and Applied Chemistry (IUPAC) met in July 1993 in order to harmonize international terminology related to detection in chemical measurement processes in analytical chemistry.[23] The following terminologies are recommended. For distinguishing a chemical amount from background noise — i.e., for making the detection decision, the critical value (L_c) of the appropriate chemical amount is used. As the measure of the inherent detection capability in chemical measurement process, the minimum detectable value (L_D) of the appropriate chemical variable is used. The terminology of "detection limit" is used as an alternative to the "minimum detectable value".[24]

The decision "detected" or "not detected" is made by comparison of the estimated concentration ($\hat{L}$) with the critical value of the respective distribution, such that the probability of exceeding L_c is no greater than 0.05 if analyte is actually absent (L = 0, null hypothesis). Thus the critical value is the minimum significant value of an estimated net concentration, applied as a discriminator against background noise. The above definition of L_c can be expressed as follows,

$$\Pr (\hat{L} > L_c \mid L = 0) \leq 0.05$$

The minimum detectable value or detection limit L_D is determined given L_c. L_D is the true concentration for which the probability that the estimated concentration $\hat{L}$ does not exceed L_c is 0.05. The above definition of L_D can be expressed as follows,[24]

$$\Pr (\hat{L} \quad L_c \mid L = L_D) \leq 0.05$$

From the above definition, L_c is more comparable to the terminology of DL in statistics than L_D. The difference is that in analytical chemistry, there is 5% probability that the non detect is larger than L_D. In contrast, in statistics, the left censored non detect is assumed to be less than or equal to DL with 100% probability. From the definition, Lc is smaller than DL in value.

Therefore, if the reported detection limit for the censored data is actually L_c, while statistic method is used to analyze censored data and L_c is taken as DL in the analysis, the influence on the results associated with censoring is actually underestimated. The reason is that a nominal smaller detection limit is used when taking the non detect x as $x \in [0, L_c]$.

A typical chemical measurement process consists of two primary substructures. The first one is sample preparation and instrumental measurement which converts the measured concentration to a signal or response.[24] The second one is an evaluation process that transforms the signal into an estimated concentration for the analyte. The first step is associated with censoring. The use of a larger volume in the sample, improvement of the sensitivity of the instrument, or both, can decrease detection limits or avoid the occurrence of censored data.

The censored data in urban air toxics emission factors are left-censored. For example, some measurements in data set for urban air toxics emission factors are reported as below a single detection limit or multiple detection limits. Multiple detection limits arise when data are collected by different sampling and analytical procedures or when data are collected with different gas sampling volumes.

In terms of the reported non detected data points and the corresponding DLs in the emission factor data, no information regarding its precise definition is available. If the measurement procedure is standard, the reported DL may be the same as L_c defined by the National Institute of Standards and Technology.[24] Thus statistical method is used to deal with censored urban air toxic emission factors, the influence by censoring is underestimated.

Commonly used methods for dealing with environmental censored data are statistically biased and are limited in their usefulness. Such methods are typically aimed only at developing an estimate of the mean of the data set. The several methods most often used include ignoring

the non-detected values, replacing non-detected values with zero, replacing non-detected values with the detection limit (DL) or replacing non-detected values with one-half of the detection limit.[25-28] These methods cause biased estimation of the mean. For example, replacing non-detected values with zero always underestimate the mean, while replacing non-detected values with DL always overestimate the mean. Ignoring the non-detected values and replacing the non-detected values with DL/2 sometimes overestimate the mean and sometimes underestimates the mean based on different data set. Furthermore, they do not provide insight regarding the population distribution and uncertainty in the mean from which the measured data are a sample. For example, the benzene emission factor from residual oil combustion was reported as 2.14×10^{-7} lb/gallon fuel combusted. It is the average of 12 data points. Two of them are detected points, and ten of them are censored points. For censored data, values equal to half of the detection limits were used to calculate the average emission factor.[29] The possible bias in this estimate and the uncertainty in the mean was not quantified.

As stated in Section 1.2, the quantification of variability and uncertainty for urban air toxics emissions is highly motivated. Since there is lack of information regarding the true value of a censored data point, variability and uncertainty associated with censoring needs to be taken into account for urban air toxics emission factors. This is different from the quantification of variability and uncertainty for typical emission factors for criteria pollutants. For criteria pollutants emission factors, since censored data issues do not commonly occur, the variability in the emission factors is reflected by the values of the detected data points. The apparent variability may include a component of random measurement error. The quantified uncertainty is explicitly caused by random sampling error, which is mainly caused by sample size. The uncertainty implicitly includes random measurement error, which is mainly influenced by inter-

unit variability.[6] For urban air toxic emission factor that contains non-detects, in order to get an unbiased estimate of the inter-unit variability, the censored data must be appropriately analyzed. The approach used here is to assume that each censored value would be between zero and its detection limit. Furthermore, the factors that influence the uncertainty estimates of the statistics and CDF are not only sample size and inter-unit variability, but also issues related to censoring. These issues include the number of non-detects and the values of the detection limits. The key motivating questions regarding the quantification of variability and uncertainty in censored data include:

1. How can the unbiased estimates of CDF and statistics (e.g. mean) for censored data be estimated?
2. How can the uncertainty in the CDF and statistics for censored data be quantified?
3. Is the uncertainty associated with censoring considerably large compared to other factors of sample size and inter-unit variability? In what situation it may be worthwhile to use a more sensitive instrument in order to reduce uncertainty associated with censoring?
4. Can the method be applied to deal with censored data outside the urban air toxic fields, such as survival data?

1.4 Probabilistic Urban Air Toxics Emission Inventories

In recent years, urban air toxics emission inventories have become available based upon point estimates of specific pollutant emissions from different source categories. One example is the 1996 Houston emission inventory. This inventory has been selected for a variety of analyses.[30] Another example is for Jacksonville, Florida, which includes 107 pollutants using 2000 as a base year.[31] The available air toxics city emission inventories are useful as a basis for development of probabilistic urban air toxic emission inventories.

The emission inventory is usually obtained by the product of emission factors and activity factors and the uncertainty in the mean emission inventories is caused by the uncertainty in the mean emission factors and activity factors. The emission factor is defined as the average emission rate from a population of sources. Because of small sample size and large inter unit variability, there is large uncertainty in the reported emission factors. In this dissertation, the quantified uncertainty in the emission inventory includes uncertainty in the mean emission factors as well as a nominal uncertainty in the mean activity factors.

The steps of the development of probabilistic emission inventory are to: (1) quantify variability and uncertainty in the urban air toxics emission factors for each source category; (2) quantify the uncertainty in urban air toxics emission inventory for each source category; and (3) quantify the uncertainty in the total emission inventory. In the first step, both uncensored data and censored data are dealt with. The motivating questions regarding the development of probabilistic urban air toxics emission inventory include:

1. How large can the uncertainty in urban air toxic emission factors for a specific source category be?
2. How large can the uncertainty in the total urban air toxic emission inventory be?
3. When there is no directly relevant data available, can surrogate data be used? Is there extra uncertainty introduced by using surrogate data?
4. When comparing results for different cities, are ranges of uncertainty similar for a given pollutant?

1.5 Objectives

Based on the backgrounds and motivations, the primary objectives of this research are:

1. To develop, test and verify a novel method to quantify variability and uncertainty in censored data;
2. To apply the method to censored urban air toxic emission factors;
3. To develop probabilistic urban air toxic emission inventories;
4. To identify the key sources of uncertainty in urban air toxic emission inventories.

1.6 Overview of Research

The research of this dissertation mainly focuses on three aspects: methodologies regarding the quantification of variability and uncertainty in censored data, applications of the methodologies, and development of probabilistic urban air toxics emission inventories.

The methodologies developed in this research include: (1) fit parametric distributions to censored data sets using MLE; (2) quantify variability and uncertainty for censored data sets using empirical bootstrap simulation; (3) test the robustness of the method in different situations with various coefficients of variation, percentages of censoring and sample sizes; (4) evaluate the reliability of MLE/Bootstrap method for estimating the mean and (5) verify the MLE/Bootstrap method with an independent method.

The application of the method includes 16 cases studies of censored urban air toxics emission factors from external combustion sources, including benzene, formaldehyde, Benzo(a)pyrene, mercury, arsenic, cadmium, total chromium, chromium VI and lead with single or multiple detection limits from coal, fuel oil and/or wood waste external combustion sources. The proportion of censored values in the emission factor data ranges from 4 to 80 percent. The key issues addressed with application of the MLE/Bootstrap method include: (1) How should inter-unit variability and uncertainty in emission factors be quantified for censored data sets? (2) What characteristics of censored data sets are important determinants of uncertainty in the mean?

(3) Are the mean estimate and the estimate of uncertainty in the mean sensitive to the choice of parametric distribution for inter-unit variability? (4) What is the relative range of uncertainty in the mean estimates of selected air toxic emission factors?

Probabilistic urban air toxics emission inventories can be developed when the variability and uncertainty in the emission factors are known. Two cities are selected to develop probabilistic urban air toxic emission inventories: Houston, TX and Jacksonville, FL. Probabilistic benzene, formaldehyde, chromium and arsenic emission inventories are developed for Houston and probabilistic 1, 3-butadiene, mercury, arsenic, benzene, formaldehyde and lead EIs are developed for Jacksonville. The variability and uncertainty in the emission factors for these pollutants are reported for different source categories. The key sources of uncertainties are identified for the emission inventories of these pollutants.

1.7 Organization

The dissertation will first present the development, test and verification of MLE/Bootstrap method for censored data, which is in Chapter 2, and then present the applications of the MLE/Bootstrap method to censored urban air toxic emission factors, which is in Chapter 3. The development of the probabilistic urban air toxics emission inventories for Houston and Jacksonville are addressed in Chapters 4 and 5, respectively. Finally, the conclusions of this study and the recommendations for future studies are presented in Chapter 6. Chapters 2 through 5 are four manuscripts that the author has submitted or plans to submit for publication in peer-reviewed journals. Each chapter of this manuscript has its own list of references cited.

1.8 References

1. Smith, R.L., French, C.L., Murphy, D.L., and Thompson, R. "Ranking and Selection of Hazardous Air Pollutants for Listing under Section 112(k) of the Clean Air Act Amendments of 1990," Technical Support Document of Integrated Urban Air Toxics Strategy, by EPA office of Air Quality Planning and Standards, 1999. http://www.epa.gov/ttn/atw/urban/urbanpg.html (accessed on 10/1/ 2000)
2. Frey, H.C., Bharvirkar, R., and Zheng, J. (1999) "Quantitative Analysis of Variability and Uncertainty in Emissions Estimation," Prepared by North Carolina State University for Office of Air Quality Planning and Standards U.S. Environmental Protection Agency, Research Triangle Park, NC.
3. Cullen, A.C. and Frey, H.C., *Probabilistic Techniques in Exposure Assessment*, Plenum Press: New York and London, 1999.
4. Frey, H.C. and Rhodes, D.S. (1996) "Characterizing, Simulating, and Analyzing Variability and Uncertainty: An Illustration of Methods Using an Air Toxics Emissions Example," *Human and Ecological Risk Assessment: an International Journal*, 2(4): 762-797.
5. Morgan, M.G. and Henrion, M. (1990), Uncertainty: A Guide to Dealing with Uncertainty in Qualitative Risk and Policy Analysis, Cambridge University Press: New York, NY.
6. Zheng, J. (2002), "Quantification of Variability and Uncertainty in Emission Estimation: General Methodology and Software Implementation," Ph.D. Dissertation, Department of Civil Engineering, North Carolina State University, Raleigh. www.lib.ncsu.edu

7. NRC, *Modeling Mobile Source Emissions*, National Academy Press: Washington D.C., 2000.
8. NRC, *Science and Judgment in Risk Assessment*, National Academy Press: Washington D.C., 1994.
9. Frey, H.C. and Zheng, J. (2002), "Probabilistic Analysis of Driving Cycle-Based Highway Vehicle Emission Factors," *Environmental Science and Technology,* 36(23): 5184-5161.
10. Frey, H.C. and Zheng, J. (2002), "Quantification of Variability and Uncertainty in Utility NO_x Emission Inventories," *Journal of Air and Waste Management Association,* 52(9): 1083-1096.
11. Frey, H.C. and Bammi, S. (2002), "Quantification of Variability and Uncertainty in Lawn and Garden Equipment NO_x and Total Hydrocarbon Emission Factors," *Journal of the Air & Waste Management Association*, 52(4): 435-449.
12. Frey, H.C. and Bammi, S. (2003), "Probabilistic Nonroad Mobile Source Emission Factors," *Journal of Environmental Engineering*, 129(2): 162-168.
13. Frey, H. C., and Bharvirkar, R. Chapter in *Risk Assessment of Environmental and Human Health Hazards: A Textbook of Case Studies*, D. Paustenbach, Ed., John Wiley and Sons: New York, 2001.
14. McCleese, D.L., and LaPuma, P.T. (2002), "Using Monte Carlo Simulation in Life Cycle Assessment for Electric and Internal Combustion Vehicles," *International Journal of Life Cycle Assessment*, 7(4): 230-236.
15. Averill, A.F., Ingram, J.M., and Nolan, P.F. (1999), "A Study of the Dispersion of Solvent Vapour in the Workspace During Wipe Cleaning of Metal Components with

Organic Solvents – A Monte Carlo Uncertainty Analysis," *Transaction of the Institute of Metal Finishing,* 77(9): 204-208.

16. Pollack, A.K.; Bhave P.; Heiken J.; Lee K.; Shepard S.; Tran C.; Yarwood G.; Sawyer R.F.; Joy B.A. *Investigation of Emission Factors in the California EMFAC7G Model*; PB99-149718INZ; Coordinating Research Council: Atlanta, GA, 1999.
17. Rao, S.T., Ku, J.Y., and Rao, K. S. (1991), "Analysis of Toxic Air Contaminant Data Containing Concentrations Below the Limit of Detection," *Journal of the Air and Waste Management Association,* 41(4): 442-448.
18. Patrick, D. et al., *Toxic Air Pollution Handbook,* Van Nostrand Reinhold: New York, 1994.
19. Miller, R. G., Gong, G., and Munoz, A. *Survival Analysis,* Wiley: New York, 1981.
20. She, N. (1997), "Analyzing Censored Water Quality Data Using A Non-parametric Approach," *Journal of the American Water Resources Association,* 33(3): 615-624.
21. Lynn, H. S. (2001), "Maximum Likelihood Inference for Left-censored HIV RNA Data," *Statistics in Medicine,* 20(1): 33-45.
22. Bogaerts, K. (2002), "Modeling Tooth Emergency Data Based On Multivariate Interval-Censored Data," *Statistics in Medicine,* 21(24): 3775-3787.
23. ISO-IUPAC, *Limits of Detection: Nomenclature Harmonization Meeting,* P.-Th. Wilrich (ISO), M. Parkany, chairman (IUPAC), L. A. Currie, Rapporteur, Washington, DC (1993). See also, Hartmann, El, "Erfassungsvermogen von Analysenverfahren," *Fresenius Z. Anal. Chem.,* 335 (1989) 954.

24. Currie, L. A. (1995), "Nomenclature in Evaluation of Analytical Methods Including Detection And Quantification Capabilities," *Pure and Applied Chemistry*, 67(10): 1699-1723.

25. Gilliom, R.J. and Helsel, D.R. (1986), "Estimation of Distributional Parameters for Censored Trace Level Water Quality Data 1. Estimation Techniques," *Water Resources Research*, 22(2): 135-146

26. Newman, M.C., Dixon, P.M., Looney, B.B., and Pinder, J.E. (1989), "Estimating Mean and Variance for Environmental Samples with Below Detection Limit Observations," *Water Resources Bulletin*, 25(4): 905-915

27. Haas, C.N. and Scheff, P.A. (1990), "Estimation of Averages in Truncated Samples," *Environmental Sciences and Technology*, 24(6): 912-919

28. Elvira, B., Rao, S., and Porter, P.S. (1999), "Identifying Pollution Source Regions using Multiple Censored Data," *Environmental Sciences and Technology*, 33(13): 2273-2277

29. EPA (1993), "Emission Factor Documentation for AP-42 Section 1.3 Fuel Oil Combustion," Prepared by Acurex Environmental Corporation, Edward Aul & Associates, Inc and E.H. Pechan and Associates, Inc, for Office of Air Quality Planning and Standards, Office of Air and Radiation, U.S. Environmental Protection Agency, Research Triangle Park, NC. http://www.epa.gov/ttn/chief/ap42/ch01/bgdocs/b01s03.pdf (accessed by 5/ 2001)

30. Hanna, S. (2003), Personal communication with Steve Hanna, Hanna Consultants, Kennebunkport, ME, via email, February 8th, 2003.

31. Tilley, L. (2003), Personal communication with Lori Tilley, Regulatory and Environmental Services Department, City of Jacksonville, FL, via email, May 26th, 2003.

Part II

QUANTIFICATION OF VARIABILITY AND UNCERTIANTY FOR CENSORED DATA SETS AND APPLICATION TO AIR TOXIC EMISSION FACTORS

Submitted to

Risk Analysis

Abstract. Many environmental data sets, such as for air toxic emission factors, contain several values reported only as below detection limit. Such data sets are referred to as "censored." Typical approaches to dealing with censored data sets include replacing censored values with arbitrary values of zero, one-half of the detection limit, or the detection limit. Here, an approach to quantification of the variability and uncertainty of censored data sets is demonstrated. Empirical bootstrap simulation is used to simulate censored bootstrap samples from the original data. Maximum Likelihood Estimation (MLE) is used to fit parametric probability distributions to each bootstrap sample, thereby specifying alternative estimates of the unknown population distribution of the censored data sets. Sampling distributions for uncertainty in statistics such as the mean, median and percentile are calculated. The robustness of the method was tested by application to different degrees of censoring, sample sizes, coefficients of variation and numbers of detection limits. Lognormal, gamma and Weibull distributions were evaluated. The reliability of using this method to estimate the mean is evaluated by averaging the best estimated means of 20 cases for small sample size of 20. The confidence intervals for distribution percentiles estimated with bootstrap/MLE method compared favorably to results obtained with the non-parametric Kaplan-Meier method. The bootstrap/MLE method is illustrated via an application to an empirical air toxic emission factor data set.

Key Words: Non-detects, Urban air toxics, Censored data sets, Maximum likelihood estimation, Bootstrap Simulation, Kaplan-Meier estimator, Monte Carlo Simulation

1.0 Introduction

The purpose of this paper is to demonstrate a methodology for quantifying variability and uncertainty in air toxic emission factors in situations in which one or more data values are not detected. Toxic air pollutants are estimated to pose human health risks in urban areas. The U.S.

Environmental Protection Agency has developed an Integrated Urban Air Toxic Strategy, which includes a framework for addressing urban air toxics emissions. EPA developed a list of 33 urban air toxics, which represent the priority for additional assessment of the health effects of air toxics in urban areas.[1]

Air toxic emissions are subject to both variability and uncertainty.[2,3] Variability refers to diversity or heterogeneity among members of population. For example, emissions of a particular pollutant typically differ from one specific emission source to another within a given source category. Uncertainty arises due to lack of knowledge regarding the true value of a quantity or regarding the true distribution for variability. Variability and uncertainty can both be represented as probability distributions in a two-dimensional modeling framework.[2, 4, 5] For example, confidence intervals can be constructed with respect to a best estimate of the cumulative distribution function (CDF) for variability. The range of the confidence intervals for the CDF represents uncertainty. Information regarding variability in urban air toxic emissions is needed to identify high emitters or highly exposed populations. Information regarding uncertainty is needed to characterize the quality of an emissions inventory and to target data collection to reduce uncertainty.

Because of inherent limitations of chemical/analytical measurement methods, emissions data sets often contain several observations reported as below a detection limit, and are usually referred to as "censored." [6] For example, data sets for emission factors of urban air toxics are often reported as censored with a single detection limit or with multiple detection limits. Multiple detection limits arise when data are collected by different sampling and analytical procedures or when data are collected with different gas sampling volumes.

Commonly used methods for dealing with environmental data sets that contain detection limits are statistically biased and are limited in their usefulness. Such methods are typically aimed only at developing an estimate of the mean of the data set. The several methods most often used include ignoring the non-detected values, replacing non-detected values with zero, replacing non-detected values with the detection limit (DL) or replacing non-detected values with one-half of the detection limit.[7, 8, 9, 10, 11] These methods cause biased estimation of the mean. Furthermore, they do not provide insight regarding the population distribution from which the measured data are a sample.

In contrast to commonly used methods, Maximum Likelihood Estimation (MLE) can be used to fit parametric distributions to censored data sets, including the portion of the distribution that is below one or more detection limits.[7, 8, 9, 10, 11] Asymptotically unbiased estimates of statistics, such as the mean, can be estimated based upon the fitted distribution. In order to estimate uncertainty in statistics, the method of bootstrap simulation is employed in this work.[12]

The objectives of this paper are: (1) to fit parametric distributions to censored data sets using MLE; (2) to quantify variability and uncertainty for censored data sets using empirical bootstrap simulation; (3) to test the robustness of the method in different situations with various coefficients of variation, percentages of censoring and sample sizes; (4) to evaluate the reliability of MLE/Bootstrap method for estimating the mean and (5) to verify the MLE/Bootstrap method with an independent method - Kaplan-Meier estimator.[13]

2.0 MLE Parameter Estimates for Censored Lognormal, Gamma and Weibull Distributions

For environmental data sets, such as concentrations or emission factors, lognormal, gamma and Weibull distributions are often chosen as parametric distributions to represent

variability in data.[2, 3, 4, 5, 14, 15] Therefore, in this paper, lognormal, gamma and Weibull distributions are used to illustrate and evaluate the MLE-based method for fitting parametric distributions to censored data.

The lognormal distribution is defined by the following probability distribution function:[16]

$$f(x \mid m, s^2) = \frac{1}{s\sqrt{2p}} \frac{e^{-(\log x - m)^2/(2s^2)}}{x} \quad (1)$$

Where $0 \le x < \infty, -\infty < m < +\infty, s > 0$; and m and s are the mean and standard deviation of ln(x). In this paper, m is defined as parameter 1 and s is defined as parameter 2.

The gamma distribution is defined by the following probability distribution function:[16]

$$f(x \mid a, b) = \frac{1}{\Gamma(a) b^a} x^{a-1} e^{-x/b} \quad (2)$$

Where $0 \le x < \infty, \; a, b > 0$. In this paper, a is defined as parameter 1 and b is defined as parameter 2.

The Weibull distribution is defined by the following probability distribution function:[17]

$$f(x|a, b) = \frac{a}{b}(x/b)^{a-1} \exp(-(x/b)^a) \quad (3)$$

Where $0 \le x < \infty$, $a, b > 0$. In this paper, a is defined as parameter 1 and b is defined as parameter 2.

The MLE technique is applicable to data sampled from various distributions and it is easily implemented in a computer program. The most general formulation of the likelihood function for censored data sets having multiple detection limits is:[18]

$$L(q_1, q_2, \ldots, q_k) = \prod_{i=1}^{n} f(x_i|q_1, q_2, \ldots, q_k)\{\prod_{m=1}^{p}(\prod_{j=1}^{ND_m} F(DL_m|q_1, q_2, \ldots, q_k))\} \quad (4)$$

Where,

x_i = Detected data point, where, i = 1, 2, …, n

$\theta_1, \theta_2, \ldots, \theta_k$ = Parameters of the distribution

ND_m = Number of non-detects corresponding to detection limit DL_m, where, m = 1, 2, …, P.

P = Number of detection limits

f = Probability density function

F = Cumulative distribution function

For data that are below detection, the cumulative probability of the detection limit is used in lieu of the likelihood.

For computational convenience, it is more common to work with the log-likelihood function instead of the likelihood function itself. According to equation (4), for the lognormal distribution, the log-likelihood function including left-censored data is given by:[18]

$$J(m,s) = -n\ln s - \frac{n}{2}\ln(2\pi) - \sum_{i=1}^{n}\left\{\frac{(\ln x_i - m)^2}{2s^2}\right\} + \sum_{m=1}^{p} ND_m \ln\left\{0.5\left[1 + erf\left(\frac{\ln DL_m - m}{s\sqrt{2}}\right)\right]\right\} \quad (5)$$

Where, erf is the error function. For the gamma distribution, the log-likelihood function including left-censored data is given by:[18]

$$J(\alpha,\beta) = -n\left\{\alpha\ln(\beta) + \ln[\Gamma(\alpha)]\right\} + \sum_{i=1}^{n}\left\{(\alpha-1)\ln(x_i) - \frac{x_i}{\beta}\right\} + \sum_{m=1}^{p} ND_m \ln\left\{\frac{\int_0^{DL_m/\beta} e^{-t}t^{\alpha-1}dt}{\Gamma(\alpha)}\right\} \quad (6)$$

For the Weibull distribution, the log-likelihoood function including left-censored data is given by:[18]

$$J(\alpha,\beta) = -n\ln\left(\frac{\alpha}{\beta}\right) + \sum_{i=1}^{n}\left\{(\alpha-1)\ln\left(\frac{x_i}{\beta}\right) - \left(\frac{x_i}{\beta}\right)^{\alpha}\right\} + \sum_{m=1}^{p} ND\ln\left\{1-\exp\left[-\left(\frac{DL}{\beta}\right)^{\alpha}\right]\right\} \quad (7)$$

Based upon earlier work by Frey and Rhodes (1996, 1998), a non-linear programming optimization algorithm was used to maximize the log-likelihood functions of Equations (5), (6) and (7).[2, 19] Computer subroutines by Press *et al.* were used for optimization and for evaluation of the various special functions required in some of the log-likelihood functions, such as the gamma function, error function, and beta function.[20]

3.0 Estimation Uncertainty in Statistics Using Bootstrap Simulation

In Parametric bootstrap simulation is widely used to estimate confidence intervals for statistics of data sets or parameters of fitted distributions in cases without censoring.[2, 3, 4, 5, 12, 14, 19] In conventional parametric bootstrap simulation, a parametric probability distribution is fit to the original data set, which has a sample size of *n* and is referred to as a "mother" distribution. Monte Carlo simulation is used to randomly simulate bootstrap samples from the "mother" distribution, each of sample size *n*. Typically, *B* bootstrap samples are simulated. For each bootstrap sample, a replication of a given statistic is calculated. For example, one can obtain *B* estimates of the mean, standard deviation, or distribution parameters. The numerical value of each replication of a statistic will differ from that of other replications because of the effect of random sampling when comparing one bootstrap sample with another. The B estimates of a statistic are used to describe a probability distribution for the statistic. A probability distribution for a statistic is also referred to as a sampling distribution. The sampling distribution is used as the basis for estimating confidence intervals. For example, the 95 percent confidence interval for the mean is based upon the interval between the 2.5^{th} and 97.5^{th} percentiles of the sampling distribution for the mean.

In the case of a censored data set, the conventional approach to parametric bootstrap simulation cannot be directly applied. Specifically, it is necessary to generate bootstrap samples so that there can be random variation in the number of data points that are below detection limit. If the data set is a single detection limit case, the data points below the detection limit can be determined as "censored". However, in a multiple-detection-limit case, a detected point may have a value larger than a detection limit of a censored value if they come from different test facilities or testing methods. In order to deal with this issue, an empirical bootstrap approach is used. In empirical bootstrap simulation, each of the original *n* data points is sampled with replacement and with equal probability of being sampled. In the original data set, either the value of data point is given for detected data or the detection limit is given for censored data. Therefore, for each data point, an indicator symbol d is given to indicate whether it is a detected value or below a detection limit. A value of d equal to 1 was used to represent a data point below a detection limit and d equal to 0 was used to represent a detected data point. In the case of non-detected data, the numerical value of the data point used in the bootstrap simulation was the detection limit itself. When generating bootstrap samples from the original censored data set, both the data point value and its indicator symbol were sampled together. Therefore, for each bootstrap sample, it is known as to which data points are detected and which data points are censored. For each bootstrap sample, MLE was used to fit a parametric distribution. Thus, B estimates of the distribution parameters and of the fitted distributions were developed. Values of the mean and standard deviation were simulated by generating 500 random samples from each of the B fitted distributions. Because the statistics of a specific percentile can not be obtained from analytical solution, numerical method is used instead of analytical method. The overall scheme of the bootstrap simulation method for censored data is described in Figure 1.

4.0 Robustness Test of the Bootstrap Method for Quantifying Uncertainty in the Statistics of Censored Data Sets

A controlled test was designed to evaluate the robustness of the MLE/Bootstrap method when applied to analyze variability and uncertainty of censored data sets with a single detection limit and with multiple detection limits. In order to do this, different population distributions were first specified for lognormal, gamma and Weibull distributions with mean equal to 1. In order to evaluate the impact of different levels of variability, for each type of parametric distribution cases were considered based upon coefficients of variation equal to 0.5, 1 and 2. In order to exactly match the assumed mean and coefficient of variation, the parameters of the assumed population distributions were determined using the Method of Matching Moments.[5, 21] The estimated parameters of each distribution are shown in Table 1. In order to evaluate different degrees of censoring with a single detection limit, different detection limits corresponding to 30% and 60% cumulative probability of censoring were obtained based upon the inverse cumulative distribution function of each assumed population distribution. Bootstrap samples were obtained from the specified population distribution. All data values in the bootstrap samples below the assumed detection limit in each case were treated as non-detects. The censoring degree here is nominal since it actually varies for each bootstrap sample. In order to evaluate the effect of different data sample sizes, three sample sizes of 20, 40 and 100 were considered for each type of distribution, coefficient of variation and degree of censoring. The MLE/Bootstrap method was applied to each specified population distribution for each type of distribution, coefficient of variation, degree of censoring and sample size. In addition to test cases with only one detection limit, cases with two detection limits and three detection limits were considered. For the tests

with multiple detection limits, 60 percent of the data set was assumed to be below one or more detection limits.

For cases with two detection limits, one detection limit was assigned to the 30 percent cumulative probability value, designated as DL2, and the other was assigned to the 60 percent cumulative probability value, designated as DL3. During bootstrap simulation, any randomly sampled value less than DL2 was treated as a non-detected value with a detection limit of DL2. Similarly, any sampled value greater than DL2 but less than DL3 was treated as a non-detected value with detection limit DL3. For cases with three detection limits, detection limits were assigned to the 10, 30 and 60 percentiles of the distribution, designated as DL1, DL2 and DL3, respectively.

For multiple detection limits as described above, non-detected values greater than one detection limit but less than another (e.g., between DL2 and DL3) were handled as interval-censored data.[22] For example, for the lognormal distribution, the log-likelihood function for interval censoring with two detection limits is:

$$J(\mu,\sigma) = -n\ln\sigma - \frac{n}{2}\ln(2\pi) - \sum_{i=1}^{n}\left\{\frac{(x_i-\mu)^2}{2\sigma^2}\right\} + ND_1\ln\left\{0.5\left[1+erf\left(\frac{DL_1-\mu}{\sigma\sqrt{2}}\right)\right]\right\} + ND_2\ln\left\{0.5\left[erf\left(\frac{DL_2-\mu}{\sigma\sqrt{2}}\right)-erf\left(\frac{DL_1}{\sigma\sqrt{2}}\right)\right]\right\} \quad (8)$$

In interval censoring, the probability that a data point is between DL1 and DL2 is equal to the cumulative probability of DL2 minus the cumulative probability of DL1. The numerical values of the detection limits for each type of distribution and coefficient of variation are given in Table 2.

The uncertainty in the cumulative distribution function of a specified parametric distribution is quantified based upon 95% confidence intervals obtained by bootstrap simulation. Figures 2 to 4 show uncertainty results for lognormal distributions with coefficients of variation of 0.5, 1 and 2, respectively. These figures show results for three sample size (10, 20 and 40) and for no censoring, 30% censoring and 60% censoring. Where applicable, the detection limit is shown by a vertical dashed line.

From Figure 2, the range of uncertainty becomes larger for the portion of the distribution below the detection limit as the percentage of censoring increases. Thus, as expected, there is more uncertainty regarding the estimate of the portion of the distribution that is below detection than there would be if all of the data had been detected. Comparing the three charts in any row of Figure 2, the 95% confidence interval of the specified distribution becomes narrower as the sample size increases from 10 to 20 to 40. Figures 3 and 4 reveals similar trends with respect to the amount of censoring and sample sizes. A comparison of Figures 2, 3 and 4 for the same percentage of censoring and the same sample sizes, reveals that the 95% confidence interval of the distribution gets wider as the coefficient of the variation increases from 0.5 to 1 to 2. Thus, the range of uncertainty is shown to increase as the range of variability increases with all else held constant. Although not shown, similar comparison were done for the gamma and Weibull distributions as a function of amount of censoring, sample size and coefficient of variation and similar results were obtained regarding the influence of the these factors on uncertainty.

In order to evaluate the influence of distribution type, sample size, coefficient of variation, and amount of censoring on the average value and 95 percent confidence interval of the mean, the MLE/Bootstrap method was conducted 10 times to each specified population distribution to get more stable results. B = 1000 were used each time. As shown in Table 3, the

estimated mean of the specified population distribution is 1.00 in each case. The MLE method does not guarantee that the moments of fitted distribution will be the same as the moments of the underlying population distribution. Nonetheless, for a sample size of 100, the average mean value from the bootstrap simulation was equal to 1.00 for the gamma and Weibull distributions for all cases considered. For the lognormal distribution and n = 100, the average mean value is 1.00 for all cases with CV = 0.5 and for the 0% and 30% censoring cases with CV = 1. The average mean value for n =100 for the lognormal tends to deviate slightly from 1.00 as the coefficient of variation increases and as the amount of censoring increased as indicated by average mean values of 1.01 to 1.03 for the cases of n = 100 and for different amounts of censoring. As sample size decreases, there tends to be more variation in the average value of the mean compared to the population mean, particularly for CV = 2.0. For example, considering all cases for CV = 0.5 and CV = 1.0, the average means range from 1.00 to 1.03, with the largest deviation occurring at 60% censoring and the smallest sample size. In contrast, for CV = 2, the average means vary from 1.00 to 1.13, with the largest deviation occurring at a sample size of 20 with 60% censoring. The gamma distribution appears to provide the most consistent estimate of the mean, with average mean values varying only from 1.00 to 1.01 over all cases considered compared to a range of 1.00 to 1.13 for the lognormal distribution and 1.00 to 1.04 for the Weibull distribution. Considering different distribution types and coefficients of variation, the worst case overall is for lognormal distribution with coefficient of variation equal to 2. Compared to other probability models, the lognormal distribution is "tail-heavy".[5] Because the results of MLE are highly influenced by values at the upper tail, the results tend to vary more for small sample size and for the lognormal distribution.

When interpreted with respect to the confidence intervals of the mean, the deviations of the average mean values compared to the population mean values are small. For example, the mean of 1.13 for the lognormal case with n = 20, CV = 2, and 60% censoring has a 95 percent confidence interval of 0.48 to 2.49. The similar case for the Weibull distribution has a mean of 1.04 and a confidence interval of 0.38 to 2.22. Both confidence intervals well enclose the population mean of 1.00. As expected, for a given distribution, CV, and amount of censoring, the confidence intervals become narrower as the sample size increases. For a given sample size and CV, the confidence intervals tend to widen as the amount of censoring increases. However, the width of the confidence interval for any given distribution, sample size, and CV is relatively insensitive to the amount of censoring. As shown in Table 2, the detection limits for 30% and 60% censoring are typically less than or approximately equal to the mean value of the distribution. However, because means are more sensitive to large values than to small values, additional uncertainty in the non-detected region of the distribution does not substantially influence the confidence interval for the mean in this case. The only apparent exception to this trend is for the lognormal distribution with n = 20 and CV = 2, for which the upper bound of the 95 percent confidence interval changes with censoring. However, of all the cases considered these are numerically the least stable and the differences in results are only modest.

The influence of multiple detection limits on the mean and 95 percent confidence interval of the mean are evaluated in Table 4. For sample sizes of 40 and 100, and particularly, for the gamma and Weibull distributions, the average mean values are typically equal to 1.00 except for a few cases, typically associated with only one detection limit, in which the average mean varies from 1.00 to 1.04. For the lognormal distribution, this type of result was obtained only for CV = 0.5. For CV = 1 and especially for CV = 2, the average means for the lognormal case range from

1.00 to 1.13. The highest average means occur for smaller sample size and fewer detection limits. A detailed investigation of the bootstrap samples revealed that the MLE method produces fitted distributions that are sensitive to sample values obtained in the upper tail. The sensitivity to such samples is greater for smaller sample sizes. Furthermore, the sensitivity to artifacts of samples from the upper tail are more influential for cases with only one detection limit versus cases with multiple detection limits. With multiple detection limits, more is known regarding the shape of the distribution in the non-detected region because the cumulative probability of each detection limit is included in the likelihood function. Thus, more constraint is imposed upon the shape of the distribution as the number of detection limits increases and, correspondingly, there is less influence of samples in the tail. Thus, although the MLE method is unbiased as sample size increases, for small sample values there can be apparent biases in the mean, especially for large CV and a tail-heavy parametric distribution. However, even in the worst case shown in Table 4, the 99 percent confidence interval of the mean well encloses the population values of 1.00.

The MLE/Bootstrap method can also be used to estimate uncertainty in statistics other than the mean, such as the median and 90^{th} percentile as illustrated in Table 5 for the example of the Weibull distribution with CV = 2 for n = 20 and n = 100. The analytical solutions for the median and 90^{th} percentile of this distribution are 0.30 and 2.67 respectively. The average estimates of the statistics are equal or very close to the analytical solution for all levels of censoring, although the difference between the simulated and population value tends to increase slightly with more censoring for n = 20. The confidence intervals for the median tend to widen slightly as the percentage of censoring increase. This is expected since the median is below the detection limit associated with 60 percent censoring. Although not shown here, the cases for the

other coefficients of variation and distribution types were also studied and similar results were obtained.

5.0 Evaluation of the Bootstrap Method for Estimating the Mean of Censored Data Sets

The objective of this section is to verify that the MLE/Bootstrap method is reliable with respect to estimation of the mean even for small sample size. For this purpose, 20 synthetic data sets with sample size 20 were randomly generated from a gamma distribution with a mean of 1 and standard deviation of 1. The samples in the synthetic data set were ranked in ascending order. Different detection limits were assigned to each synthetic data set in order to achieve different amounts of censoring including 30% and 60%. That is, for each synthetic data, the detection limit corresponding to 30% censoring is assigned a value equal to the average of the 6^{th} and 7^{th} ranked data points in the 20 data points, and the detection limit corresponding to 60% censoring is assigned a value equal to the average of the 12^{th} and 13^{th} ranked data points. The MLE/Bootstrap method was applied to get the best estimate of the mean. The t-test was applied to the 20 best estimates of means to test if their average was statistically significantly different from the population mean of 1. The results are shown in Table 6.

For all three levels considered in Table 6, the p-values of the t test are larger than the significance level of 0.05. Therefore, the hypothesis can not be rejected which indicates that the best estimate of the means from the MLE/Bootstrap method are reliable for synthetic data sets.

In order to compare the best estimate of the mean from the MLE/Bootstrap method with that from conventional methods, 5 methods were applied to the sample data sets as summarized in Table 7. For 0% censoring, the averages of the best estimates of the means from all of the conventional methods are the same and equal to the arithmetic average of the original data sets.

The MLE/Bootstrap result of 0.990 is slightly different than the arithmetic average because MLE does not preserve the moments. However, the estimated mean is within one percent of the population mean.

The methods of ignoring non-detected values and averaging only detected data, and of replacing non-detected values with the detection limit, are shown to overestimate the mean. The bias becomes more pronounced as the amount of censoring increases. Conversely, the method in which non-detected values are assigned a value of zero clearly underestimates the mean. The commonly used method of replacing non-detected values with one-half of the detection limit appears to be unbiased for the example of 30% censoring but the bias inherent in this method becomes obvious at 60% censoring. In the latter case, the population mean is overestimated by 6.5%. In contrast, the MLE/Bootstrap method provides consistent estimates of the mean regardless of the amount of censoring. The maximum deviation of the estimated mean from the population mean in these three cases of censoring is only 1.4 percent. Although not shown in the paper, the same tests for lognormal and gamma distributions were also done and similar conclusion can be drawn.

6.0 Verification of the Uncertainty Estimates from the Bootstrap Method with the Kaplan-Meier Estimator Method

The purpose of this section is to verify the accuracy of the novel method-MLE/Bootstrap method and the coded program for estimating the 95% confidence interval of the CDF when the population distribution is known. The confidence intervals for the CDF estimated from the MLE/Bootstrap method were compared with the results from the Kaplan-Meier estimator method. The Kaplan-Meier estimator is a non-parametric method which is used in analyzing right-censored data in survival analysis. However, with a simple transformation, the Kaplan-

Meier estimator can be used in analyzing left-censored data sets.[13] For this purpose, synthetic data sets with sample size of 40 were randomly generated from lognormal, gamma and Weibull distributions with mean equal to 1 and coefficients of variation equal to 0.5, 1 and 2. The samples in the randomly generated data set were ranked in ascending order. Different detection limits were assigned to the synthetic data sets in order to achieve 0%, 30% and 60% censoring. The MLE/Bootstrap method was applied to each synthetic data to get the CDF and the 95% confidence interval of the CDF.

The test is to compare the 95% confidence interval of the CDF estimated by MLE/Bootstrap method and the Kaplan-Meier estimator. Since the Kaplan-Meier estimator method is a non-parametric method, 40 data points were evenly drawn from the CDF estimated by the MLE method and an empirical CDF and its 95% confidence interval were calculated for these 40 points using the Kaplan-Meier estimator method. Figures 5 to 7 show the comparison of the results of the confidence intervals of the CDF estimated from MLE/bootstrap and the Kaplan-Meier estimator methods for lognormal, gamma and Weibull distributions, respectively.

Both MLE and the Kaplan-Meier estimator are asymptotically unbiased methods. However, for finite sample size, the results of the estimated CDFs from both methods are different since their basic theories are different. The MLE is the parameter point for which the observed sample is most likely.[16] The Kaplan-Meier estimator is based upon a conditional probability that a measurement after transformation will be greater than the i^{th} ordered observation given that it is greater than the $(i-1)^{th}$ ordered observation used to estimate the survival function.[13] The 95% confidence intervals estimated from the two methods are not exactly the same but they are similar. For example, for the lognormal distribution, when CV = 1 and 30 percent censoring, the largest discrepancy of the results for the 95% confidence interval

of the CDF from these two methods is only 2.7 percent for the lower level and 2.9 percent for the upper level. Similar inferences can be made in other cases shown in the figures.

7.0 Case Study for Estimation of Variability and Uncertainty in an Emission Factor Based Upon Censored Data

As an example of an emission factor data set that contains censoring, a data set for formaldehyde emissions from external combustion sources fired with bituminous and subbituminous coal combustion was selected. The data set is given in Table 8. The data set contains 14 data points, of which 5 are censored. Each of the five censored data points has a different detection limit. There are 2 data values that are greater than the largest of the five detection limits and 2 that are smaller than the smallest of the five detection limits. There are three detected data values that are greater than the smallest detection limit but smaller than the second smallest detection limit. There is one detected data value between the second and third detection limits. There is one detected data value between the fourth and fifth detection limits.

When replacing censored data with DL/2, the lognormal distribution is the best fit for the data set based on Kolmogorov - Smirnov test at significance level of 0.05. As a screening step, lognormal distribution model was used to fit to the formaldehyde emission factor data set using MLE. It is shown as a dashed white line in Figure 8. In Figure 8, since there is uncertainty regarding what value of cumulative probability to assign to 7 of the detected data points that are less than the largest detection limit, the seven data points have possible cumulative probabilities. For example, the true but unknown values of the five non-detected data points could all be less than the numerical value of the smallest detection limit, which would imply that the smallest detected data point could have a rank of 6. However, it is also possible that the smallest detected data value could have a rank of 1 if the true but unknown values of all nondetected data points

are larger. The uncertainty regarding the possible cumulative probabilities of the detected data points is indicated by solid vertical lines in Figure 8. For the 2 values that are larger than the largest detection limit, there is no uncertainty regarding the cumulative probability in these cases. The five detection limits are indicated by dashed vertical line.

From Figure 8, the fitted distribution agrees reasonably well with the observed data. The 95 percent confidence interval on the fitted distribution encloses all of the detected values. The graphical comparison of the data points and fitted distribution and its confidence intervals further confirm that lognormal distribution is a reasonable model to represent the inter-unit variability in the formaldehyde emission factors.

The detected formaldehyde emission factor data vary over almost two orders-of-magnitude and the fitted distribution spans approximately three orders-of-magnitude. Therefore, there is a large amount of uncertainty in the mean value associated with the small sample size, large amount of variability in the data, and the presence of non-detected measurements in the data set. The mean estimated based upon the fitted lognormal distribution was found to be 1.8×10^{-5} lb formaldehyde emitted per ton of coal combusted. The 95 percent confidence interval for the mean ranges from minus 77 percent to plus 209 percent of the mean value. The asymmetry of this confidence interval is based upon the large amount of variability in the data, the relatively small sample size, and the fact that an emission factor must be non-negative. The mean could be estimated from the conventional methods as well. The estimated means are 2.7×10^{-5}, 1.7×10^{-5}, 2.0×10^{-5} and 2.2×10^{-5} lb formaldehyde emitted per ton of coal combusted when using the methods of removing non-detects, replacing non-detects with zero, replacing non-detects with DL/2 and replacing non-detects with DL, respectively. The estimated mean from MLE/Bootstrap method is slightly larger than the results from replacing non-detects with zero, and smaller than

the result based upon removing the non-detects, replacing non-detects with DL/2 or DL in this case.

8.0 Conclusions

MLE is a flexible and reasonably robust method for fitting parametric distributions to censored data, based upon extensive case studies that address lognormal, gamma and Weibull distributions, sample sizes of 20 to 100, coefficients of variation of 0.5 to 2, and differing amounts and types of censoring, including single and multiple detection limits involving as much as 60 percent of the distribution. The capability of bootstrap simulation to estimate uncertainty in statistics of censored data sets represented by MLE fitted distributions was demonstrated for the cumulative distribution function, mean, median, and 90^{th} percentile. The uncertainty estimate for a statistic is not substantially sensitive to censoring as long as the numerical value of the statistic is comparable to or greater than the largest detection limit. However, it is clear that the ranges of uncertainty of the portions of the CDF below the detection limit become larger with more censoring. Statistics that are sensitive to large values of a data set, such as the mean, may not be particularly sensitive to uncertainty associated with left-censoring. The MLE/Bootstrap method was shown to be asymptotically unbiased with respect to the mean and with respect to the 95 percent confidence interval for the CDF in the non-censored region. Especially for larger sample size, the MLE/Bootstrap method was shown to give consistently unbiased estimates of the mean in comparison to conventional methods in which non detects are ignored or assigned arbitrary values of zero, one half of the detection limit, or the detection limit. The methodology for application of the MLE/Bootstrap method for dealing with censoring was illustrated by application to an empirical environmental data set for formaldehyde emissions.

The key advantages of the MLE/Bootstrap method is that it is a statistically rigorous, robust, and asymptotically unbiased method that can be used to make inferences for a wide variety of situations, including different types of distributions, coefficients of variation, sample sizes, and amounts of censoring with either single or multiple detection limits. Compared to simplified conventional methods, which are biased, the MLE/Bootstrap method is more computationally intensive. However, unlike conventional methods, the MLE/Bootstrap method enables estimation of uncertainty for any statistic, including the influence of uncertainty associated with censoring itself. The MLE/Bootstrap method can be incorporated as part of a two dimensional framework in which variability and uncertainty are distinguished. The MLE/Bootstrap method is recommended for consideration in environmental, risk assessment, and other policy relevant analyses in which censoring of data are present.

9.0 Acknowledgements

The authors acknowledge the support of the Science to Achieve Results (STAR) grants program of the U.S. Environmental Protection Agency, which funded this work under Grant No. R826790. Although the research described in this paper has been funded wholly or in part by the U.S. EPA, this paper has not been subject to any EPA review and therefore does not necessarily reflect the views of the Agency, and no official endorsement should be inferred.

10.0 References

1. Smith, R.L., French, C.L., Murphy, D.L., and Thompson, R. “Ranking and Selection of Hazardous Air Pollutants for Listing under Section 112(k) of the Clean Air Act Amendments of 1990,” Technical Support Document of Integrated Urban Air Toxics

Strategy, by EPA office of Air Quality Planning and Standards, 1999. http://www.epa.gov/ttn/atw/urban/urbanpg.html (accessed by 10/1/ 2000)

2. Frey, H.C. and Rhodes, D.S. (1996) "Characterizing, Simulating, and Analyzing Variability and Uncertainty: An Illustration of Methods Using an Air Toxics Emissions Example," *Human and Ecological Risk Assessment: an International Journal,* 2(4): 762-797.
3. Frey, H.C., Bharvirkar, R., and Zheng, J. (1999) "Quantitative Analysis of Variability and Uncertainty in Emissions Estimation," Prepared by North Carolina State University for Office of Air Quality Planning and Standards U.S. Environmental Protection Agency, Research Triangle Park, NC.
4. Frey, H.C. and Zheng, J., (2002) "Quantification of Variability and Uncertainty in Utility NO_x Emission Inventories," *Journal of Air and Waste Management Association,* 52(9): 1083-1096
5. Cullen, A.C. and Frey, H.C., (1999) *Probabilistic Techniques in Exposure Assessment*. Plenum Press: New York and London
6. Rao, S.T., Ku, J.Y., and Rao, K. S., (1991) "Analysis of Toxic Air Contaminant Data Containing Concentrations Below the Limit of Detection," *Journal of the Air and Waste Management Association,* 41(4): 442-448
7. Gilliom, R.J. and Helsel, D.R. (1986), "Estimation of Distributional Parameters for Censored Trace Level Water Quality Data 1. Estimation Techniques," *Water Resources Research,* 22(2): 135-146

8. Newman, M.C., Dixon, P.M., Looney, B.B., and Pinder, J.E. (1989), "Estimating Mean and Variance for Environmental Samples with Below Detection Limit Observations," *Water Resources Bulletin*, 25(4): 905-915
9. Haas, C.N. and Scheff, P.A. (1990), "Estimation of Averages in Truncated Samples," *Environmental Sciences and Technology*, 24(6): 912-919
10. Elvira, B., Rao, S., and Porter, P.S. (1999), "Identifying Pollution Source Regions using Multiple Censored Data," *Environmental Sciences and Technology*, 33(13): 2273-2277
11. Clarke, J.U. (1994), "Evaluating methods for statistical analysis of less than detection limit data using simulated small samples. 2. General Results," *International Conference on Dredging and Dredged Material Placement*, vol. 1: 747-755
12. Efron, B. and Tibshirani, R.J. (1993), *An Introduction to the Bootstrap,* Chapman & Hall: New York.
13. She, N. (1997), "Analyzing Censored Water Quality Data Using A Non-parametric Approach," *Journal of the American Water Resources Association,* 33(3): 615-624.
14. Frey, H.C. and Bammi, S. (2002), "Quantification of Variability and Uncertainty in Lawn and Garden Equipment NO_x and Total Hydrocarbon Emission Factors," *Journal of the Air & Waste Management Association*, 52(4): 435-449
15. Seinfeld, J.H. (1986), *Atmospheric Chemistry and Physics of Air Pollution,* John Willey and Sons, New York
16. Casella, G. and Berger, R.L. (2001) *Statistical Inference*. Duxbury Press: Pacific Grove, CA.
17. Morgan, M.G. and Henrion, M. (1990) Uncertainty: A Guide to Dealing with Uncertainty in Qualitative Risk and Policy Analysis, Cambridge University Press: New York, NY.

18. Cohen, A.C. and Whitten, B. (1988), Parameter Estimation in Reliability and Life Span Models, M. Dekker: New York.

19. Frey, H.C. and Rhodes, D.S. (1998) "Characterization and simulation of uncertainty frequency distributions: Effects of Distribution Choice, Variability, Uncertainty, and Parameter Dependence," *Human and Ecological Risk Assessment*, 4(2):423-468.

20. Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P. (1992). *Numerical Recipes in Fortran: The Art of Scientific Computing, 2nd ed.,* Cambridge University Press: New York.

21. Frey, H.C. and Zheng, J. (2002) "Technical Documentation of the AuvTool Software Tool for Analysis of Variability and Uncertainty," Prepared by Department of Civil Engineering, North Carolina State University for Office of Research and Development, U.S. Environment Protection Agency, Research Triangle Park, NC.

22. Samuelsen, S.O. and Kongerud, J. (1994) "Interval Censoring in Longitudinal Data of Respiratory Symptoms in Aluminium Potroom Workers: A Comparison of Methods," *Statistics in Medicine*, 13(17): 1771-1780.

Table 1. Parameters of the Assumed Lognormal, Gamma and Weibull Population Distributions with Coefficients of Variation of 0.5, 1 and 2

Distribution	CV[a]	Parameter1[b]	Parameter2[b]
Lognormal	0.5	-0.112	0.472
	1	-0.347	0.833
	2	-0.804	1.269
Gamma	0.5	4	0.25
	1	1	1
	2	0.25	4
Weibull	0.5	2.101	1.129
	1	1	1
	2	0.543	0.575

[a] Coefficient of Variation

[b] Refer to equations (1), (2) and (3) for parameter definitions

Table 2. Detection Limits Associated with Different Cumulative Probabilities for Specified Population Distributions as a Function of Distribution Type and Coefficient of Variation

Distribution	CV[a]	Cumulative Probability for Detection Limit		
		10%	30%	60%
Lognormal	0.5	0.488	0.698	1.008
	1	0.243	0.457	0.873
	2	0.088	0.230	0.617
Gamma	0.5	0.436	0.691	1.044
	1	0.105	0.357	0.916
	2	0.00027	0.022	0.377
Weibull	0.5	0.387	0.691	1.083
	1	0.105	0.357	0.916
	2	0.0091	0.0861	0.490

Table 3. Average and 95 Percent Confidence Intervals of Mean Estimated from Bootstrap Simulation for Lognormal, Gamma and Weibull Distribution, Sample Sizes of 20, 40 and 100, Coefficients of Variation of 0.5, 1 and 2, and Censoring of 0, 30and 60 Percent[a]

CV[b]	CP[c]	Mean	Lognormal Distribution			Gamma Distribution			Weibull Distribution		
			N = 20	N = 40	N = 100	N = 20	N = 40	N = 100	N = 20	N = 40	N = 100
0.5	0%	Average	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
		2.5 percentile	0.80	0.85	0.90	0.79	0.85	0.90	0.79	0.85	0.89
		97.5 percentile	1.24	1.17	1.11	1.23	1.17	1.11	1.23	1.16	1.11
	30%	Average	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
		2.5 percentile	0.79	0.84	0.89	0.78	0.84	0.90	0.78	0.84	0.89
		97.5 percentile	1.24	1.17	1.11	1.23	1.17	1.11	1.23	1.16	1.11
	60%	Average	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**
		2.5 percentile	0.76	0.83	0.89	0.70	0.80	0.87	0.71	0.81	0.87
		97.5 percentile	1.25	1.17	1.11	1.25	1.18	1.12	1.25	1.18	1.12
1.0	0%	Average	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
		2.5 percentile	0.64	0.73	0.81	0.61	0.71	0.80	0.61	0.71	0.80
		97.5 percentile	1.52	1.36	1.23	1.51	1.36	1.23	1.50	1.35	1.23
	30%	Average	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
		2.5 percentile	0.65	0.73	0.81	0.61	0.71	0.80	0.61	0.71	0.80
		97.5 percentile	1.55	1.37	1.23	1.50	1.35	1.23	1.50	1.34	1.23
	60%	Average	**1.03**	**1.02**	**1.02**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**
		2.5 percentile	0.63	0.72	0.82	0.56	0.68	0.78	0.58	0.69	0.79
		97.5 percentile	1.58	1.39	1.26	1.51	1.36	1.23	1.51	1.35	1.23
2.0	0%	Average	**1.04**	**1.01**	**1.01**	**1.00**	**1.00**	**1.00**	**1.02**	**1.01**	**1.00**
		2.5 percentile	0.47	0.58	0.69	0.32	0.48	0.62	0.37	0.51	0.64
		97.5 percentile	2.13	1.73	1.45	2.08	1.73	1.47	2.15	1.77	1.48
	30%	Average	**1.08**	**1.04**	**1.03**	**1.00**	**1.00**	**1.00**	**1.03**	**1.01**	**1.00**
		2.5 percentile	0.47	0.58	0.70	0.32	0.47	0.62	0.37	0.50	0.64
		97.5 percentile	2.38	1.85	1.52	2.05	1.75	1.48	2.17	1.79	1.49
	60%	Average	**1.13**	**1.05**	**1.03**	**1.01**	**1.00**	**1.00**	**1.04**	**1.01**	**1.00**
		2.5 percentile	0.48	0.58	0.70	0.34	0.47	0.62	0.38	0.51	0.64
		97.5 percentile	2.49	1.89	1.53	2.07	1.73	1.48	2.22	1.80	1.50

Table 3. Continued

[a] The method was applied for 10 times, each with B = 1000. The average of the results for 10 times was reported.

[b] CV = Coefficient of Variation, the population mean of each distribution is 1.00.

[c] CP = Censoring Percentage

Table 4. Average and 95 Percent Confidence Intervals of Mean Estimated from Bootstrap Simulation for Lognormal, Gamma and Weibull Distributions, Sample Sizes of 20, 40 and 100, Coefficients of Variation of 0.5, 1 and 2, and Multiple Detection Limits.[a]

CV[b]	DLs[c]	Mean	Lognormal Distribution			Gamma Distribution			Weibull Distribution		
			N = 20	N = 40	N = 100	N = 20	N = 40	N = 100	N = 20	N = 40	N = 100
0.5	1	Average	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**
		2.5 percentile	0.76	0.83	0.89	0.70	0.80	0.87	0.71	0.81	0.87
		97.5 percentile	1.25	1.17	1.11	1.25	1.18	1.12	1.25	1.18	1.12
	2	Average	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
		2.5 percentile	0.79	0.84	0.89	0.78	0.84	0.90	0.77	0.84	0.89
		97.5 percentile	1.25	1.17	1.11	1.23	1.17	1.11	1.24	1.16	1.11
	3	Average	**0.99**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
		2.5 percentile	0.79	0.84	0.90	0.78	0.84	0.90	0.78	0.84	0.89
		97.5 percentile	1.23	1.17	1.11	1.23	1.17	1.11	1.23	1.16	1.11
1.0	1	Average	**1.03**	**1.02**	**1.02**	**1.00**	**1.00**	**1.00**	**1.01**	**1.00**	**1.00**
		2.5 percentile	0.63	0.72	0.82	0.56	0.68	0.78	0.58	0.69	0.79
		97.5 percentile	1.58	1.39	1.26	1.51	1.36	1.23	1.51	1.35	1.23
	2	Average	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
		2.5 percentile	0.64	0.73	0.81	0.60	0.70	0.80	0.60	0.70	0.80
		97.5 percentile	1.55	1.37	1.23	1.51	1.34	1.23	1.50	1.35	1.23
	3	Average	**1.01**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**
		2.5 percentile	0.64	0.73	0.81	0.60	0.70	0.80	0.61	0.71	0.80
		97.5 percentile	1.55	1.36	1.23	1.49	1.35	1.24	1.50	1.35	1.23
2.0	1	Average	**1.13**	**1.05**	**1.03**	**1.01**	**1.00**	**1.00**	**1.04**	**1.01**	**1.00**
		2.5 percentile	0.48	0.58	0.69	0.34	0.47	0.62	0.38	0.51	0.64
		97.5 percentile	2.49	1.89	1.53	2.07	1.73	1.48	2.22	1.80	1.50
	2	Average	**1.08**	**1.04**	**1.03**	**1.00**	**1.00**	**1.00**	**1.03**	**1.01**	**1.00**
		2.5 percentile	0.47	0.58	0.70	0.32	0.47	0.63	0.37	0.50	0.64
		97.5 percentile	2.38	1.85	1.52	2.07	1.75	1.48	2.17	1.79	1.50
	3	Average	**1.05**	**1.02**	**1.01**	**1.00**	**1.00**	**1.00**	**1.02**	**1.01**	**1.00**
		2.5 percentile	0.46	0.57	0.69	0.32	0.47	0.62	0.36	0.50	0.65
		97.5 percentile	2.24	1.76	1.46	2.09	1.75	1.46	2.15	1.77	1.48

Table 4. Continued
[a] The method was applied for 10 times, each with B = 1000. The average of the results for 10 times was reported.
[b] CV = Coefficient of Variation, the population mean of each distribution is 1.00.
[c] CP = Censoring Percentage

Table 5. Average and 95 Percent Confidence Intervals of the Median and 90th Percentile Estimated from Bootstrap Simulation for the Weibull Distribution, Sample Sizes of 20 and 100, Coefficient of Variation = 2, and 0%, 30% and 60% Censoringa

Censoring Percentage	Statistics	N = 20		N = 100	
		Median	90^{th} Percentile	Median	90^{th} percentile
0%	Average	0.33	2.62	0.30	2.62
	2.5 percentile	0.11	0.94	0.18	1.67
	97.5 percentile	0.72	5.50	0.46	3.90
30%	Average	0.33	2.62	0.30	2.61
	2.5 percentile	0.089	0.94	0.17	1.66
	97.5 percentile	0.75	5.55	0.47	3.93
60%	Average	0.35	2.58	0.30	2.61
	2.5 percentile	0.055	0.91	0.15	1.65
	97.5 percentile	0.81	5.50	0.50	3.91

[a] The method was applied for 10 times, each with B = 1000. The average of the results for 10 times was reported.

Table 6. Test of Hypothesis that the Average of Bootstrap Means Are Not Statistically Significantly Different from the Population Mean: Gamma Distribution, Population Mean = 1, Population Coefficient of Variation = 1, Sample Size = 20, Three Levels of Censoring

Case[a]	Best Estimate of Mean[b]		
	0% Censoring	30% Censoring	60% Censoring
1	0.986	0.991	1.010
2	0.996	1.000	1.160
3	1.039	1.022	0.995
4	1.005	1.025	1.034
5	0.963	0.951	1.009
6	1.031	1.035	1.101
7	0.992	0.976	0.938
8	1.000	0.980	0.939
9	0.970	0.977	1.026
10	0.946	0.946	0.967
11	0.995	1.016	0.962
12	0.956	0.940	0.903
13	1.032	1.022	0.983
14	1.024	1.008	0.977
15	0.955	0.954	0.979
16	1.013	1.012	1.125
17	0.964	0.948	1.024
18	0.964	0.957	0.918
19	0.945	0.942	0.932
20	1.027	1.016	1.033
Average	0.990	0.986	1.001
Standard Deviation	0.0309	0.0325	0.0678
Min	0.945	0.940	0.903
Max	1.039	1.035	1.160
Hypothesis	The average of the mean for the 20 cases = 1		
p-value for T-test	0.1701	0.0675	0.9610
Conclusion	Hypothesis Not Rejected		

[a] Synthetic data sets with sample size 20, generated from population distribution of Gamma (1,1)

[b] Average value of the mean based upon bootstrap simulation with B = 500

Table 7. Comparison of the Average Mean Value Based Upon 20 Synthetic Data Sets from a Gamma Distribution with Mean = 1, Coefficient of Variation = 1, and Sample Size = 20 for Three Levels of Censoring and Five Estimation Methodsa

Estimation Methods	Percentage of Censoring		
	0%	30%	60%
Ignoring non-detected values	0.995	1.351	1.921
Replacing values below DL with zero	0.995	0.946	0.768
Replacing values below DL with DL/2	0.995	0.996	1.065
Replacing values below DL by DL	0.995	1.046	1.181
MLE/Bootstrap method[b]	0.990	0.986	1.001

[a] average on the best estimated means for the same data sets as in Table 6 with different methods

[b] based upon the average of bootstrap simulation for 20 data sets, each with B = 500

Table 8. Input Data of Formaldehyde Emission Factors

Rank	Value (10^{-4} lb/ton)	Indicator Symbol	Note
1	0.173	0	Detected
2	0.224	0	Detected
3	0.298	1	Censored, the value is DL
4	0.346	0	Detected
5	0.357	0	Detected
6	0.607	0	Detected
7	0.614	1	Censored, the value is DL
8	0.95	0	Detected
9	1.3	1	Censored, the value is DL
10	1.41	1	Censored, the value is DL
11	2.22	0	Detected
12	3.08	1	Censored, the value is DL
13	3.78	0	Detected
14	15.6	0	Detected

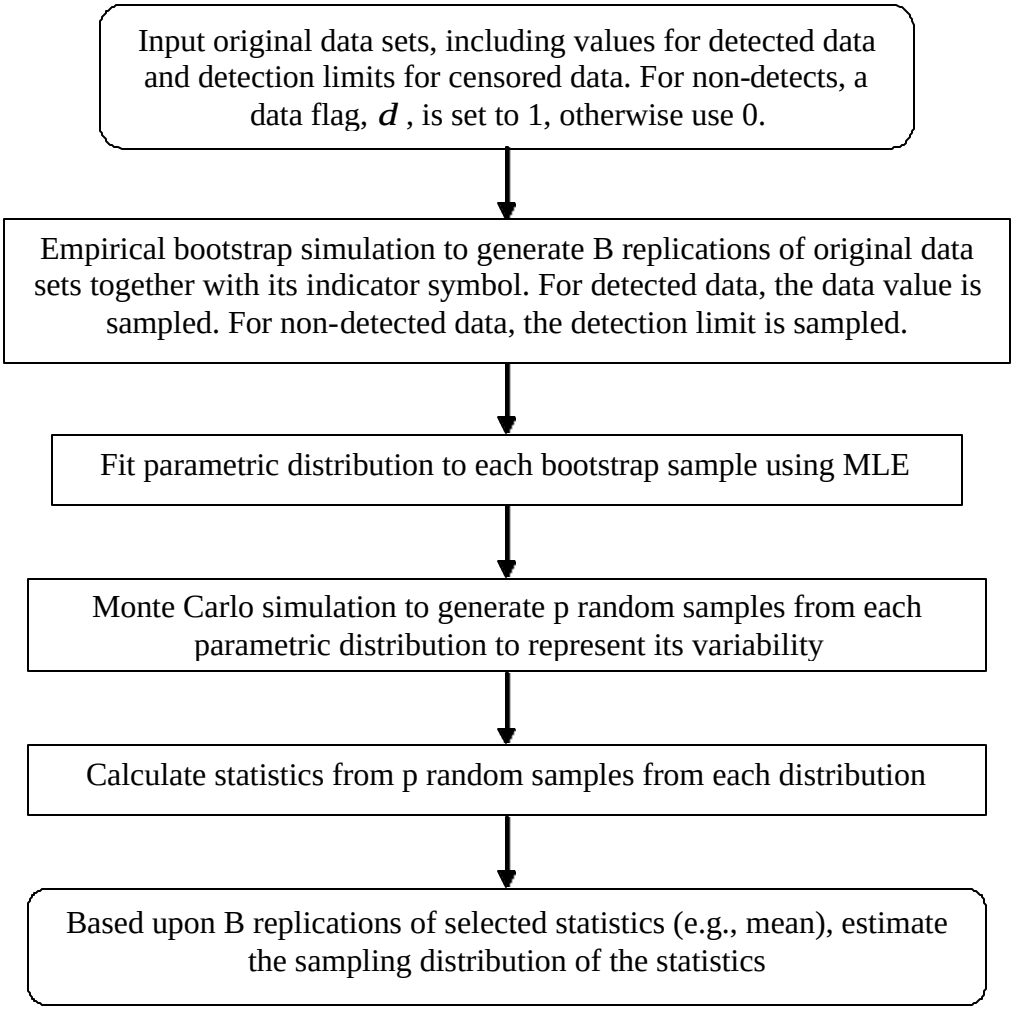


Figure 1. Scheme of Quantification of Variability and Uncertainty for Censored Data Sets

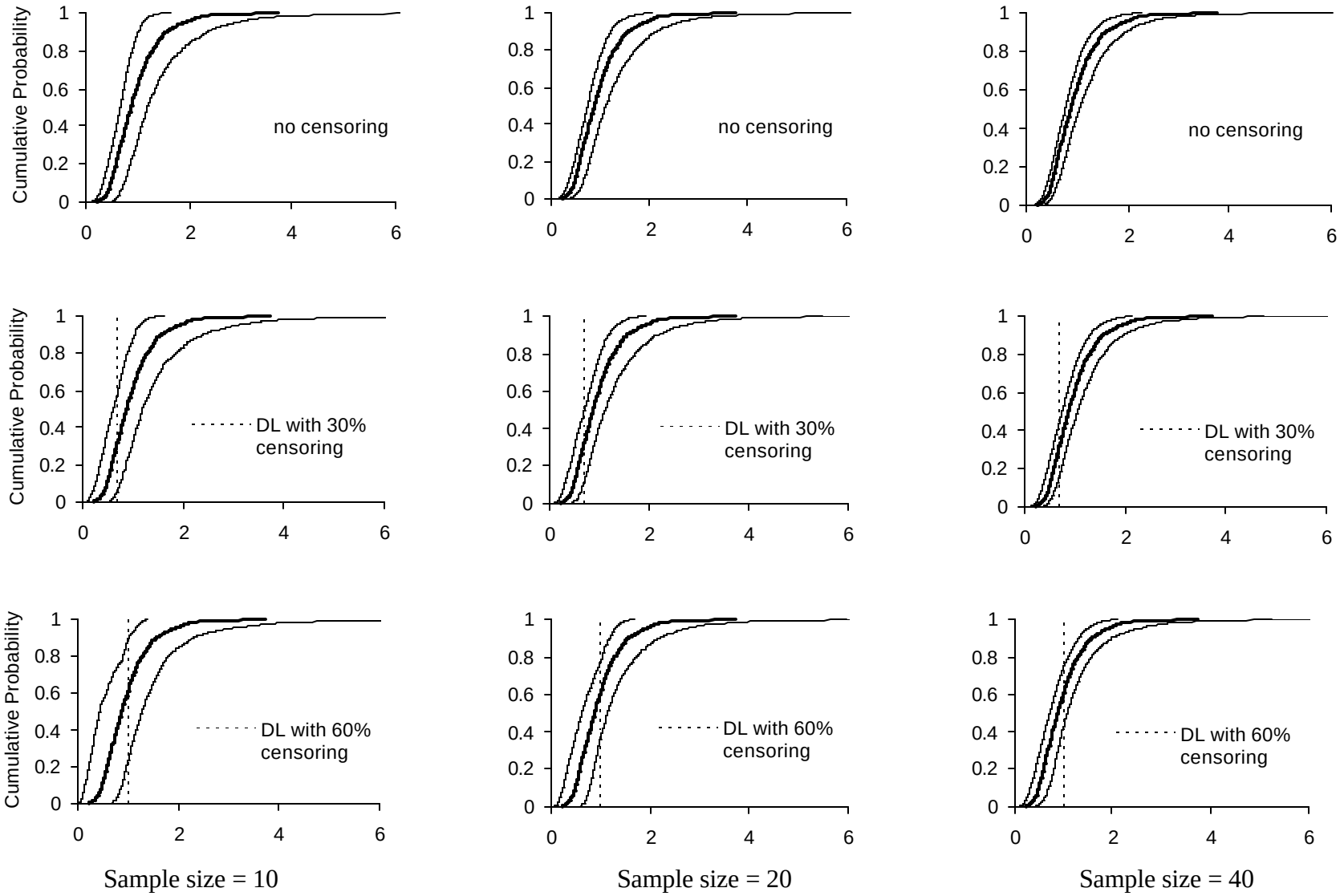


Figure 2. 95 Percent Confidence Intervals for Uncertainty in the Cumulative Distribution Function for Lognormal Distribution with Coefficient of Variation of 0.5, Sample Sizes of 10, 20 and 40, With No Censoring, 30% Censoring, and 60% Censoring (B=500)

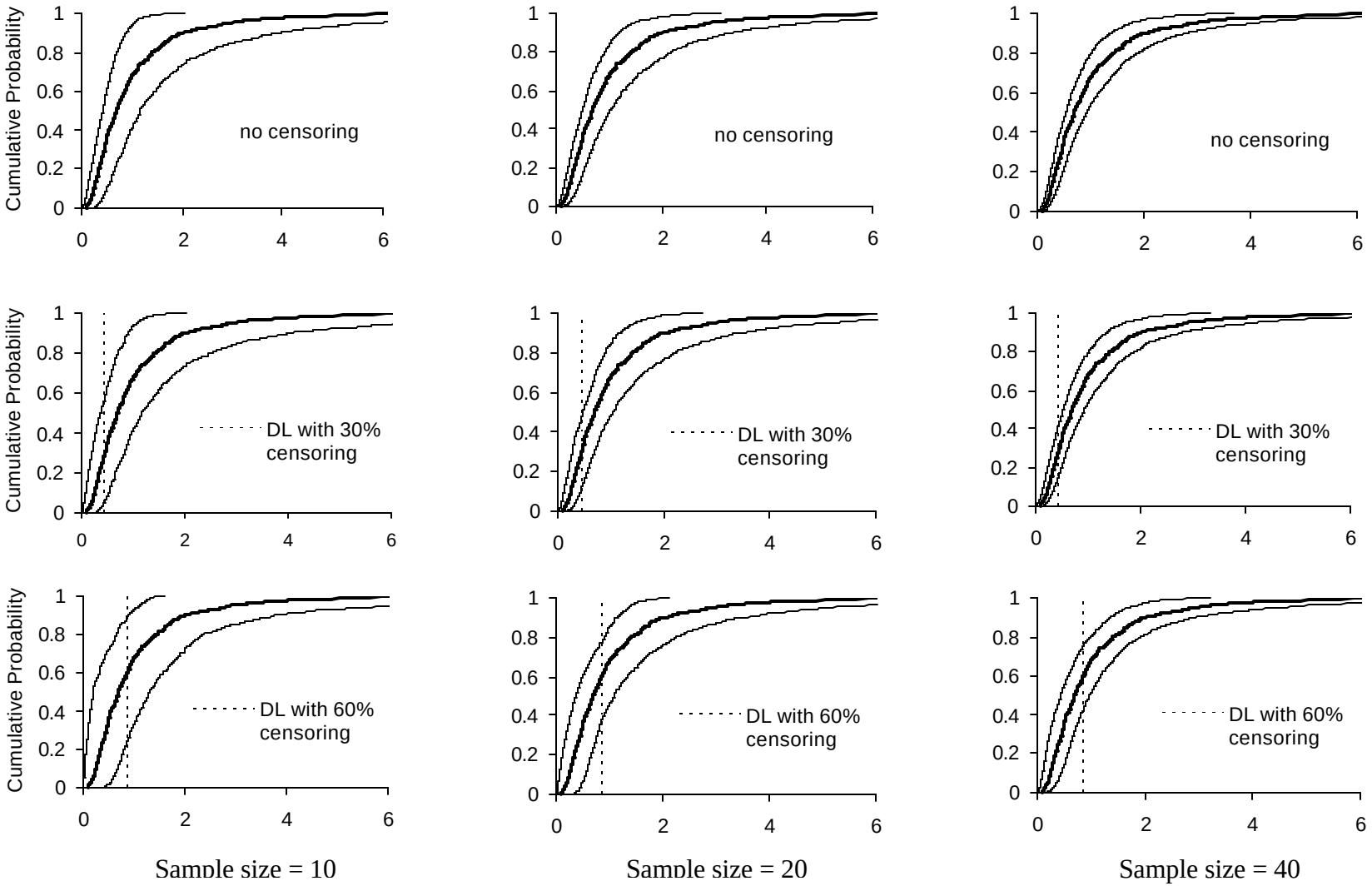


Figure 3. 95 Percent Confidence Intervals for Uncertainty in the Cumulative Distribution Function for Lognormal Distribution with Coefficient of Variation of 1, Sample Sizes of 10, 20 and 40, With No Censoring, 30% Censoring, and 60% Censoring (B=500)

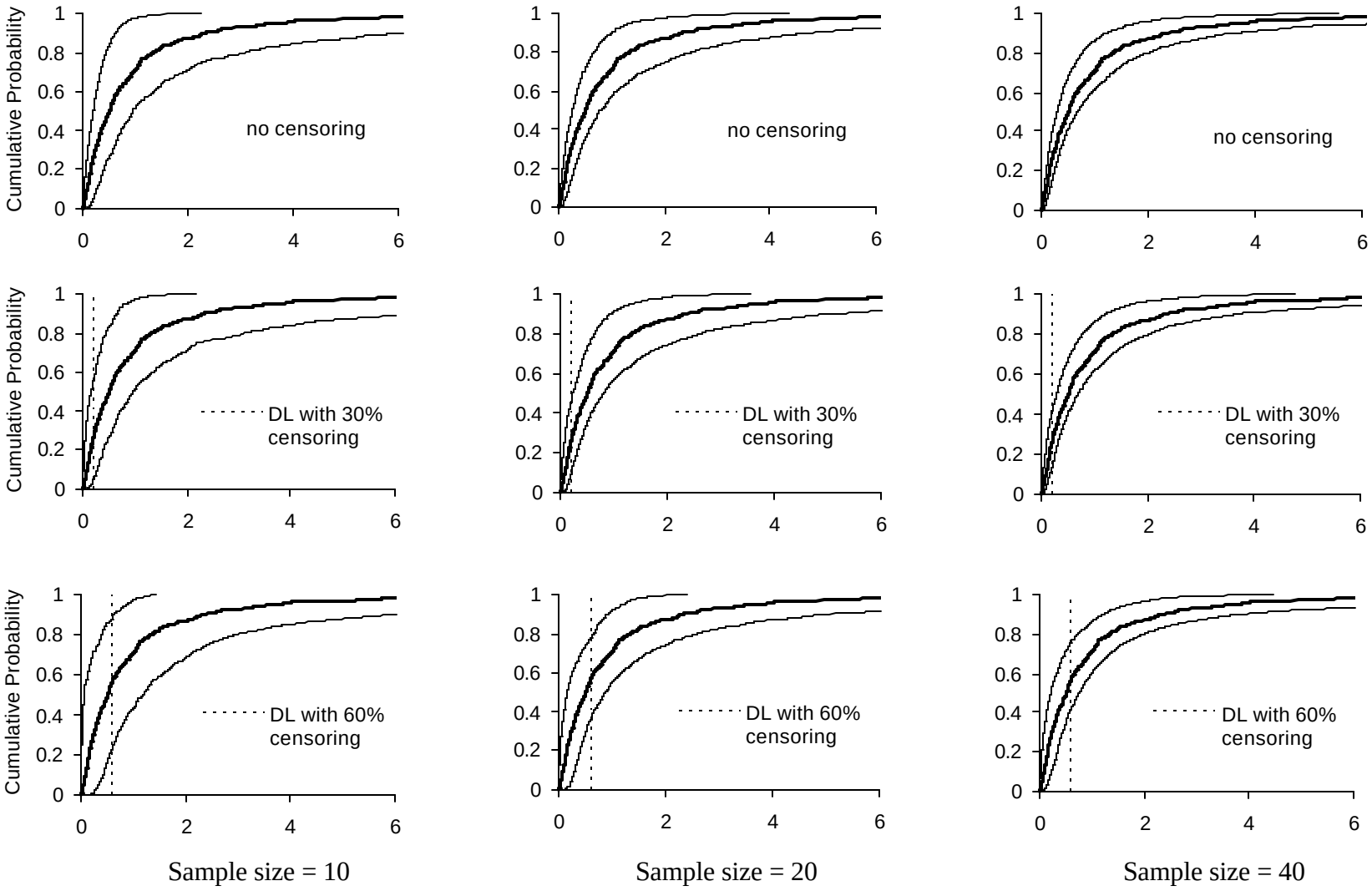


Figure 4. 95 Percent Confidence Intervals for Uncertainty in the Cumulative Distribution Function for Lognormal Distribution with Coefficient of Variation of 2, Sample Sizes of 10, 20 and 40, With No Censoring, 30% Censoring, and 60% Censoring (B=500)

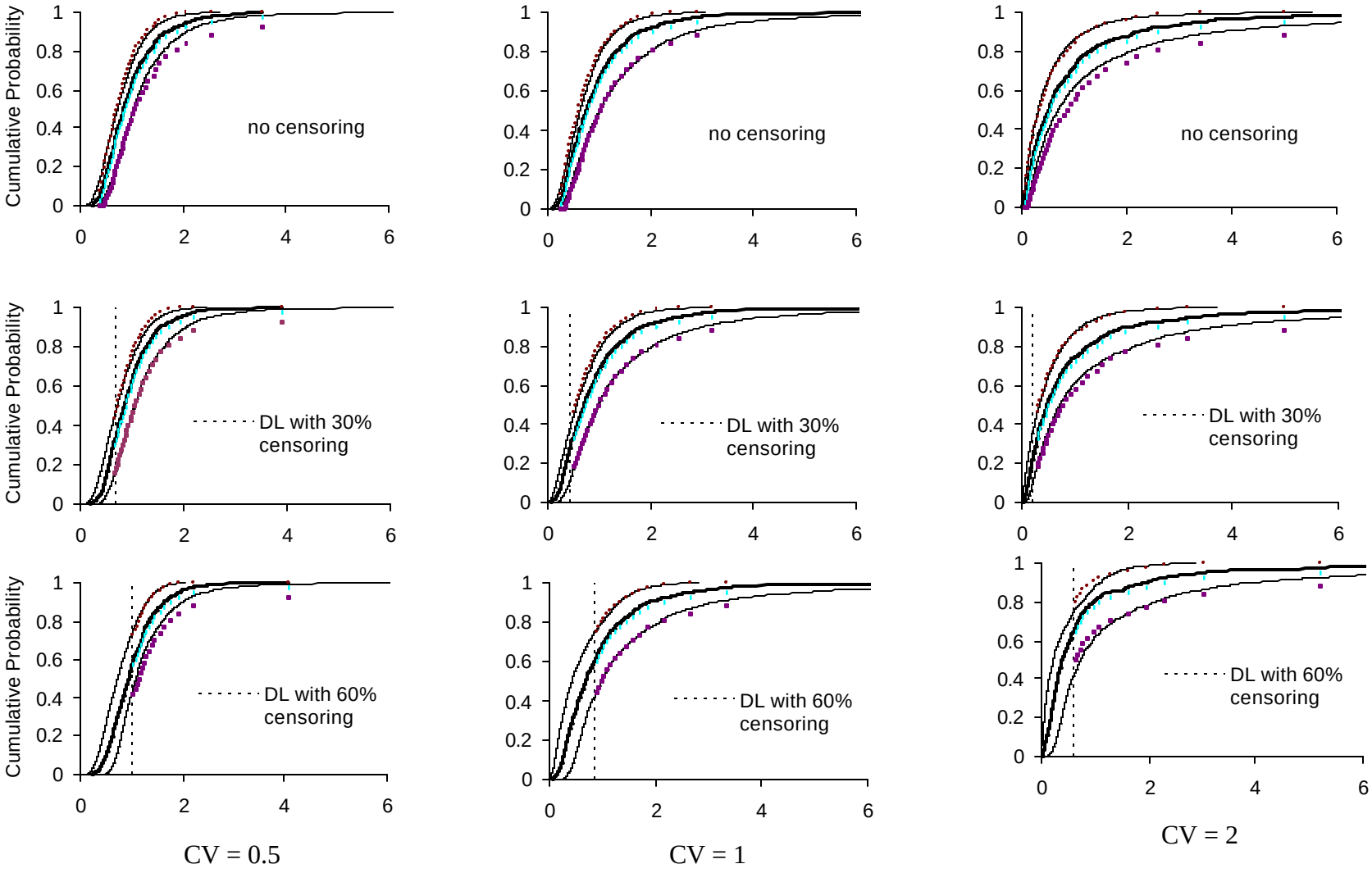


Figure 5. Comparison of 95 Percent Confidence Intervals of Cumulative Distribution Functions for the MLE/Bootstrap and Kaplan-Meier Estimator Methods for Different Amounts of Censoring and Coefficients of Variation for Lognormal Distributions (Results from MLE/Bootstrap Method Are Represented by Continuous Lines and Results from Kaplan-Meier Estimator Are Represented by Dots, Sample Size = 40)

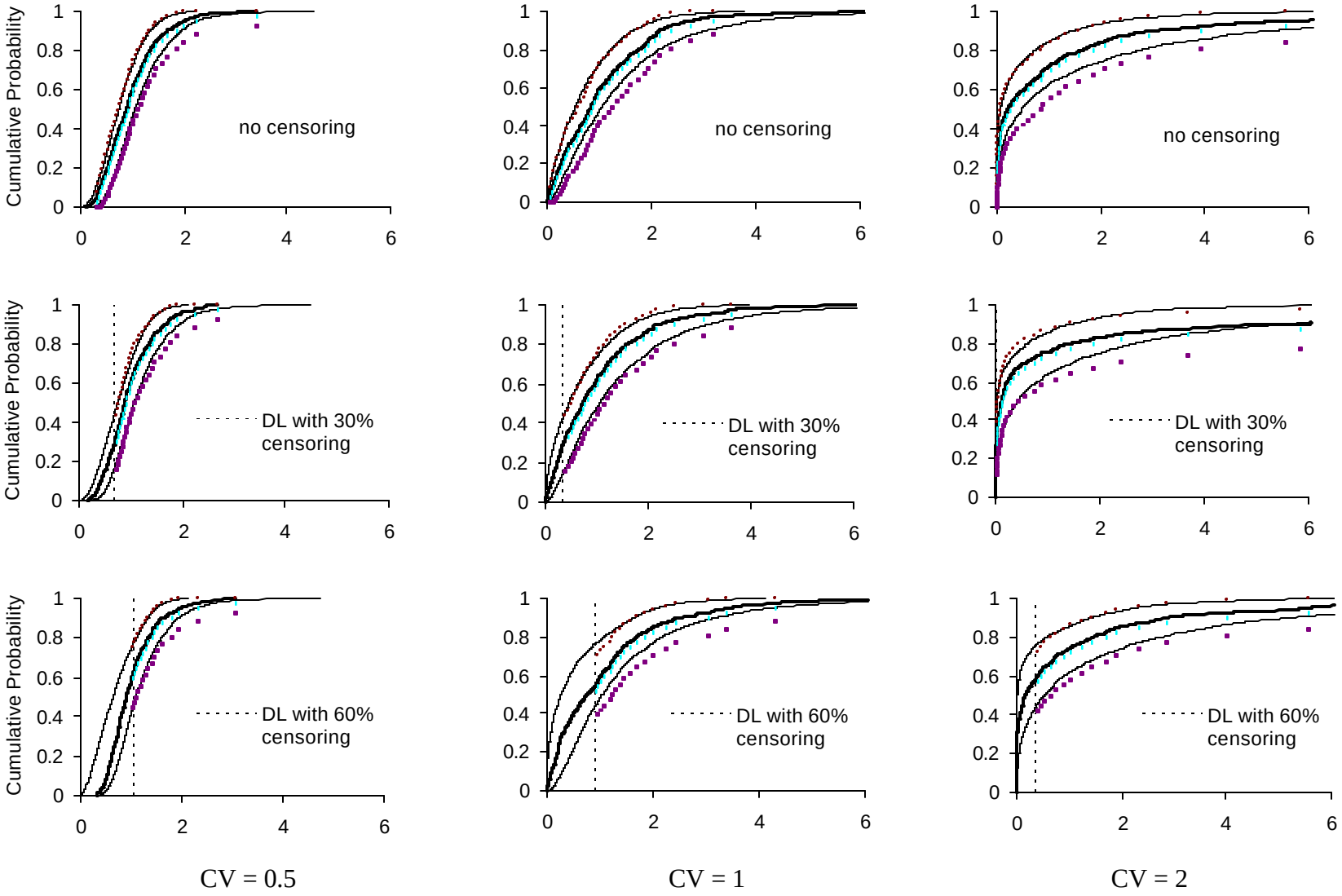


Figure 6. Comparison of 95 Percent Confidence Intervals of Cumulative Distribution Functions for the MLE/Bootstrap and Kaplan-Meier Estimator Methods for Different Amounts of Censoring and Coefficients of Variation for Gamma Distributions (Results from MLE/Bootstrap Method Are Represented by Continuous Lines and Results from Kaplan-Meier Estimator Are Represented by Dots, Sample Size = 40)

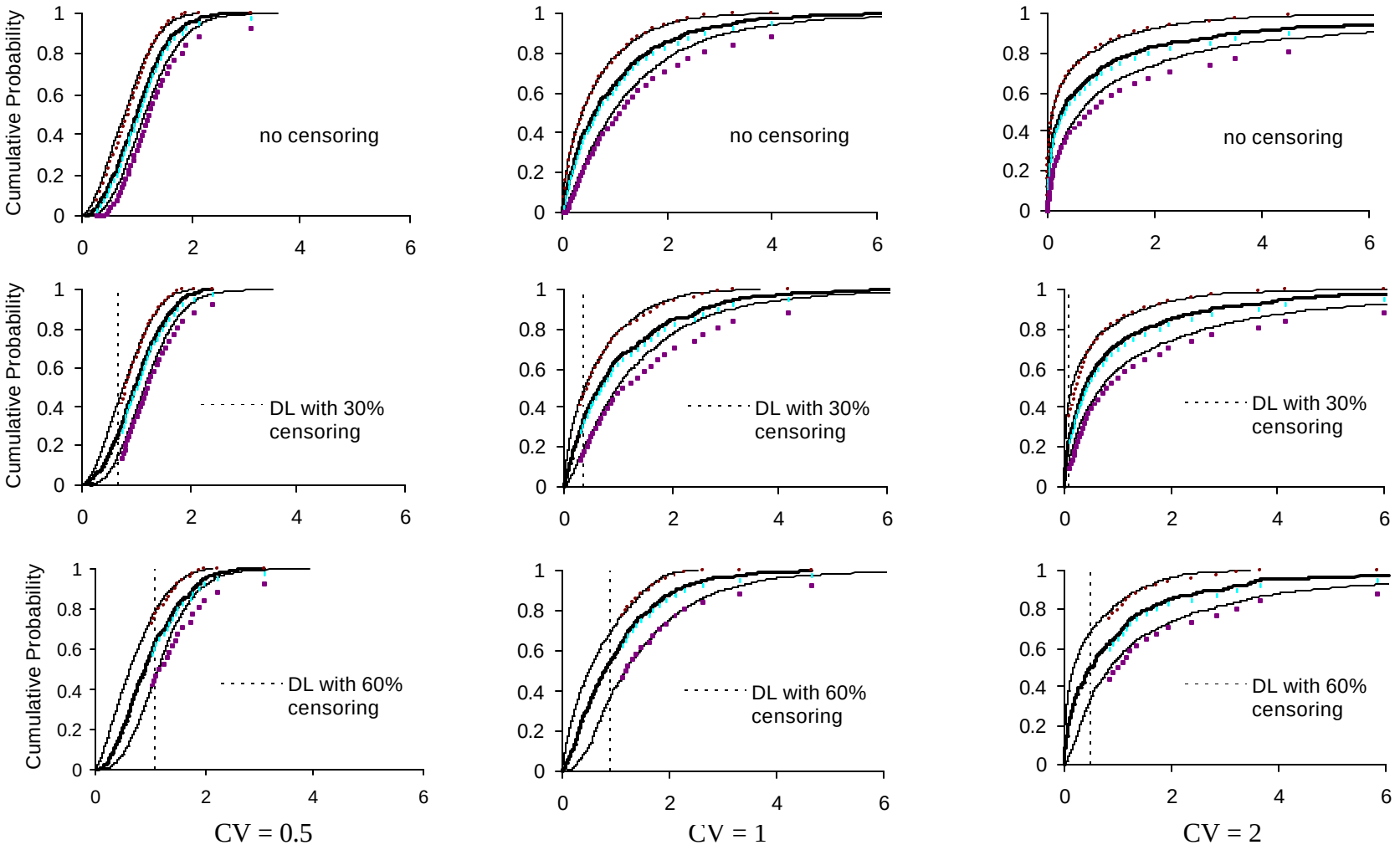


Figure 7. Comparison of 95 Percent Confidence Intervals of Cumulative Distribution Functions for the MLE/Bootstrap and Kaplan-Meier Estimator Methods for Different Amounts of Censoring and Coefficients of Variation for Weibull Distributions (Results from MLE/Bootstrap Method Are Represented by Continuous Lines and Results from Kaplan-Meier Estimator Are Represented by Dots, Sample Size = 40)

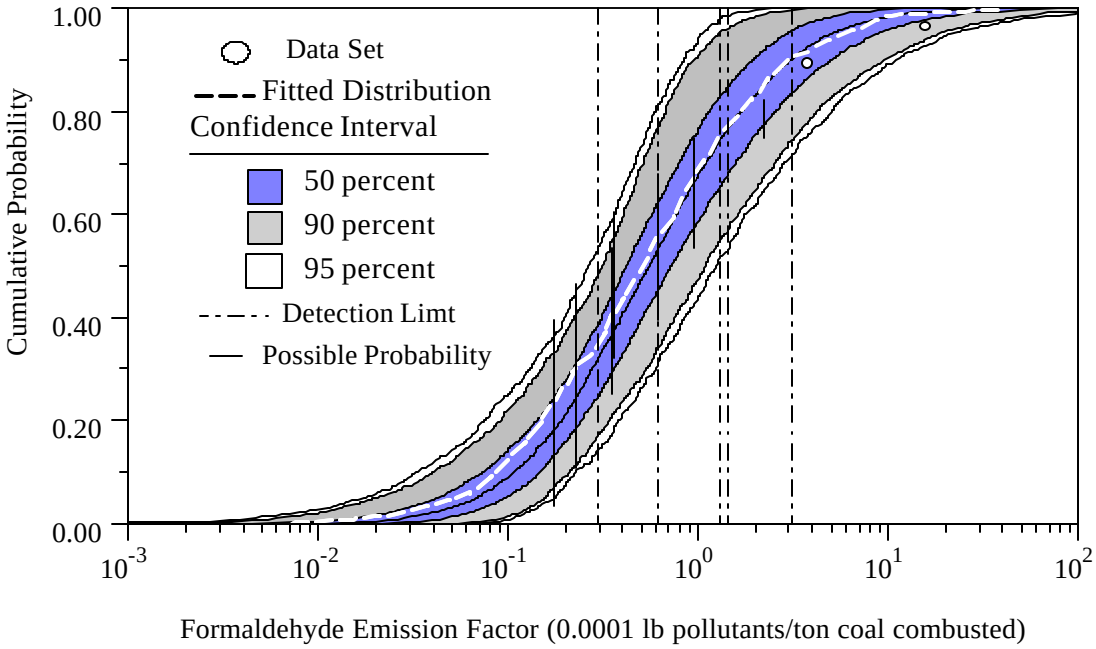


Figure 8. Variability and Uncertainty in the Formaldehyde Emission Factor for a Combustion Source Estimated Based Upon a Lognormal

Part III

UNCERTAINTY FOR DATA WITH NON-DETECTS: AIR TOXIC EMISSIONS FROM COMBUSTION

Prepared to be submitted to

Journal of Environmental Engineering

Abstract. Air toxic emission factor data often contain one or more censored points below a single or multiple detection limits. Conventional methods used to deal with censored data sets include removing non-detects, or replacing the censored points with zero, half of the detection limit or the detection limit. However, the estimated means of the censored data set by conventional methods are usually biased. Maximum likelihood estimation (MLE) and bootstrap simulation have been demonstrated as a statistically robust method to quantify variability and uncertainty of censored data set and can provide asymptotically unbiased mean estimates. Here, the MLE/bootstrap method is applied to 16 cases of censored air toxic emission factors, including benzene, formaldehyde, Benzo(a)pyrene, mercury, arsenic, cadmium, total chromium, chromium VI and lead with single or multiple detection limits from coal, fuel oil and/or wood waste external combustion sources. The proportion of censored values in the emission factor data ranges from 4 to 80 percent. The largest range of uncertainty in the mean was obtained for the external coal combustion benzene emission factor, with a 95 percent probability range of minus 93 to plus 411 percent of the mean.

Key Words: Urban air toxics, Emission factor, Censored data sets, Maximum likelihood estimation, bootstrap simulation

1.0 Introduction

The U.S. Environmental Protection Agency (EPA) has developed a list of 33 urban air toxics that represent the priority for additional assessment of the health effects of human exposure to air toxics in urban areas (Smith et al. 1999). Emission estimates are one of the key inputs to an exposure assessment, and are the focus of this paper. Because exposure assessments require explicit quantification of variability and uncertainty (NAS, 1994; EPA, 1997; Cullen and

Frey, 1999), there is a need for methods and case studies or probabilistic emission estimates for air toxics.

A key challenge to quantification of inter-unit variability in emissions is that, because of inherent limitations of chemical/analytical measurement methods, emission factor data sets of urban air toxics often contain several observations reported as below a detection limit (DL), and are referred to as "censored." These data sets can have multiple detection limits. Multiple detection limits arise when individual measurements are collected by different sampling and analytical procedures at different facilities within a source category. An ability to accurately estimate average emissions, and the variability and uncertainty in emissions, based upon censored data is a key need in urban air toxics exposure and risk assessment. If non-detected measurements are not properly accounted for, exposure and risk could be significantly misestimated (Patrick 1994).

One of the important emission source categories for urban air toxics are external combustion sources. These sources include steam-electric generating plants, industrial boilers, and commercial and residential combustion systems, such as for space heating. External combustion sources, using fuels including coal, fuel oil, and wood waste, has been estimated to contribute 70 percent of mercury emissions and 85 percent of arsenic emissions in the case of Jacksonville, FL (Tilley 2003), and are considered to be significant emission sources of these and other pollutants in other urban areas. Emission factor data for these sources are available via EPA publications (EPA 1993a, b, 2001). These data include proportions of censoring ranging from 4 to 80 percent. Thus, there is a need to develop accurate estimates of the mean emission factors and their uncertainty.

Zhao and Frey (2003) demonstrated an asymptotically unbiased method based upon maximum likelihood estimation (MLE) and bootstrap simulation for quantifying inter-unit variability, and uncertainty in statistics such as the mean, for censored data sets. In contrast, this paper demonstrates the use of the MLE/bootstrap method for estimation of variability and uncertainty for an extensive set of emission factors. For each of the three fuel sources considered, the available emission factor data include one or more of benzene, formaldehyde, mercury, arsenic, cadmium, chromium, lead, and Benzo(a)pyrene (B(a)p), for a total of 16 emission factor case studies.

1.1 Sources of Variability and Uncertainty

Variability refers to the heterogeneity across different elements of a population over time or space. Uncertainty is lack of knowledge regarding the true value of a quantity. Uncertainty is typically attributable to: (1) random measurement errors (lack of precision); (2) systematic errors (bias or lack of "accuracy") such as would be caused by imprecise calibration or use of surrogate data; (3) lack of empirical basis such as would occur when measurements have not been taken or when estimating emissions for a future source; or (4) human error such as random mistakes in entering or processing data. Variability and uncertainty can both be represented as probability distributions in a two-dimensional modeling framework. For example, confidence intervals can be constructed with respect to a best estimate of the cumulative distribution function (CDF) for variability. The range of the confidence intervals for the CDF represents uncertainty (Bogen and Spear 1987; Frey and Rhode 1996; Webster and Shih 1996; Cullen and Frey 1999; Frey and Zheng 2002a, b; Frey and Bammi 2002; Sadiq et al 2002).

1.2 Variability and Uncertainty in Urban Air Toxic Emissions

Air toxics emissions are subject to both variability and uncertainty (Patric 1994; Frey and Rhode 1996; Frey and Bharvirkar 2002). Variability and uncertainty in air toxics emissions are a contributing factor to variability and uncertainty in estimates of exposure and risk. Quantification of variability and uncertainty in air toxics emissions is needed to identify high emitters or highly exposed populations as well as to characterize the quality of an emissions inventory and to target data collection to reduce uncertainty.

The objective of this paper is to demonstrate the use of the MLE/bootstrap method to quantify the variability and uncertainty in censored air toxics emission factor data, based upon case studies for 16 censored emission factor data sets for external combustion sources. The key questions addressed in this paper include:

- How should inter-unit variability and uncertainty in emission factors be quantified for censored data sets?
- Is the mean estimate and the estimate of uncertainty in the mean sensitive to the choice of parametric distribution for inter-unit variability?
- What characteristics of censored data sets are important determinants of uncertainty in the mean?
- What is the relative range of uncertainty in the mean estimates of selected air toxic emission factors?

2.0 Methodology

Conventional approaches for estimating the mean of censored data sets are briefly reviewed. The key elements of the MLE/bootstrap simulation, and the approaches used to assess goodness-of-fit of parametric distributions compared to the empirical data, are presented.

2.1 Conventional Approaches to Dealing with Censored Data

Hass and Scheff (1990) have compared four conventional methods for estimating means of censored data:

- Using values only above DL to calculate a mean value, and ignoring information regarding non-detects, which leads to a biased estimate of the mean;
- Replacing values below DL with zero, which leads to an underestimate of the true mean;
- Replacing values below DL by DL/2, which leads to an approximate but biased estimate of the true mean;
- Replacing values below DL by DL, which leads to an overestimate of the true mean;

The conventional methods produced biased estimates of the mean and of other statistics, such as the variance. The bias typically worsens as the amount of censoring increases. In contrast, bias-corrected MLE estimates were found to be more accurate compared to other methods (Newman et al 1989; Gilliom and Helsel 1986; Elvira et al 1999; Burmaster and Wilson 2000; Frey and Zhao 2003). Moreover, an approach based upon fitting distributions to data using MLE enables more accurate insight regarding the entire distribution of variability than can be obtained via the four conventional approaches (Zhao and Frey 2003). Bootstrap simulation can be applied to estimate uncertainty in any statistic of the fitted distribution.

2.2 Method of Maximum Likelihood Estimation

In order to fit a parametric distribution representing inter-unit variability to censored data, MLE is used to estimate the distribution parameters based upon the observed sample of data. When applying MLE to left-censored data, the cumulative probability of the detection limit is used in lieu of the likelihood for each non detected measurement. The likelihood function for censored data sets having multiple detection limits is (Zhao and Frey 2003):

$$L(\theta_1, \theta_2, \ldots \theta_3) = \prod_{i=1}^{n} f(x_i | \theta_1, \theta_2, \ldots, \theta_k) \left\{ \prod_{m=1}^{p} \left(\prod_{j=1}^{ND_m} F(DL_m | \theta_1, \theta_2, \ldots, \theta_k) \right) \right\} \quad (2)$$

Where,

$\theta, \theta_2, \ldots, \theta_k$ = Parameters of the distribution

x_i = Detected data point, where, i = 1, 2, …, n

ND_m = Number of non-detects corresponding to detection limit DL_m, where, m = 1, 2, …, P.

P = Number of detection limits

f( ) = Probability density function

F( ) = Cumulative distribution function

2.3 Lognormal, Gamma and Weibull Distributions

For environmental data sets, such as concentrations or emission factors, lognormal, gamma and Weibull distributions are often chosen as parametric distributions to represent variability in data (Seinfeld, 1986; Frey and Rhode 1996; Cullen and Frey 1999; Frey and Zheng 2002a, b; Frey and Bharvirkar 2002; Frey and Bammi 2002). One of the most widely used distributional forms in probabilistic assessment is the lognormal distribution. The lognormal distribution describes random variables resulting from multiplicative processes (Ott 1990; Ott 1995). The gamma distribution is non-negative, positively skewed, and similar to the lognormal distribution in many cases but it is less "tail heavy" (Cullen and Frey 1999). The Weibull distribution is a flexible distribution that can assume negatively skewed, symmetric, or positively skewed shapes (Cullen and Frey 1999). It also may be used to represent non-negative quantities. In this paper, these three distribution types are used to analyze censored air toxics emission factor data. The fitted parametric distribution is used to represent inter-unit variability.

2.4 Bootstrap Simulation

Parametric bootstrap simulation is often used to estimate confidence intervals for statistics of data sets or parameters of fitted distributions in cases without censoring (Efron and Tibshirani 1993; Frey and Rhodes 1996; Cullen and Frey 1999; Frey and Bammi 2002; Frey and Zheng 2002a; Sadiq et al 2002; Frey and Bharvirkar, 2002; Zhao and Frey 2003; Faraggi, 2003). In conventional parametric bootstrap simulation, a parametric probability distribution representing variability is fit to the observed data, which has a sample size of n. To simulate random sampling error, Monte Carlo simulation is used to randomly simulate B synthetic data sets, referred to as bootstrap samples, each of sample size n. In the approach used here, parametric distributions are fit to each bootstrap sample. The sampling distribution of a given statistic, such as the mean, variance, or distribution percentiles are estimated based upon B replications of the statistic. For censored data set with only one detection limit that is smaller than all observed value in the data, parametric bootstrap simulation can also be used. In each bootstrap sample, simulated values below the detection limit are identified as non-detects. Thus, these bootstrap samples would be composed of censored data sets with the same detection limit as the original data set.

However, in the case of censored data set with multiple detection limits, or with a single detection limit that is larger than at least one detected point, the parametric bootstrap simulation method discussed above can not be directly applied. In order to generate bootstrap samples with random variation in the number of non-detects and for which some detection limits may be larger than the sample values of some observations, empirical bootstrap simulation using bootstrap pairs is used. Each data point in the original data set is paired with a binary indicator variable. The binary indicator variable denotes whether the data point is an observation or a detection

limit. Both the data point and the indicator symbol are sampled together in an empirical bootstrap simulation. Therefore, the status of each data point is known. The MLE method for fitting distributions to censored data is applied to each bootstrap sample to get *B* replications of the CDF and of statistics of interest (Zhao and Frey 2003). In this way, the variability and uncertainty of multiply censored data can be quantified.

2.5 Evaluation of Goodness-of-fit

The Kolmogorov-Smirnov (K-S) test and graphical comparison of the CDF of the fitted distribution to the data are widely used to evaluate the goodness-of-fit of parametric distributions fit to uncensored data (Morgan and Henrion 1990; Cullen and Frey 1999; Lu, 2003). However, the K-S test cannot be directly applied to in the case of a censored data set. To gain semi-quantitative insight regarding goodness-of-fit for a censored data set, an approximation procedure is used. For this purpose only, each non-detected measurement was replaced with one half of its detection limit to create a modified data set. The K-S test was applied to a distribution that was fit to the modified data set. If the fitted distribution was not rejected the K-S test at significance level of 0.05 for the modified data, it was taken as a reasonable candidate parametric distribution for the original censored data. The candidate type of parametric distribution was then fit to the censored data taking into account the presence of non-detects, resulting in different parameter estimates than for the modified data set. This is a semi-quantitative approach that can be used to guide the selection of several candidate probability distribution models to fit to censored data. A comparison of the bootstrap confidence intervals for the CDF of the fitted distribution to the observed data was applied to confirm the adequacy of the fit and to help guide the choice regarding a preferred distribution. The larger the proportion of data contained within the confidence intervals, the stronger the preference toward a particular candidate distribution.

3.0 Case Studies

In this section, the variability and uncertainty of 16 censored air toxic emission factor data sets were quantified. The results regarding selection of a best fitting parametric distribution for inter-unit variability, the MLE parameter estimates of the fitted distribution, and the bootstrap confidence intervals for the mean, are given. The mean estimates were compared with those from conventional methods.

3.1 Data

Empirical censored air toxic emission factor data from external combustion sources were obtained from background documents of EPA's report AP-42 that compiles emission factors for stationary sources (EPA 1993a, b; EPA 2001). Urban air toxic pollutants for which data were available for one or more of the three fuels considered include benzene, formaldehyde, benzo(a)pyrene (B(a)p), mercury, arsenic, cadmium, chromium and lead.

Table 1 summarizes the available emission factor data with respect to the pollutant, fuel, sample size, percentage of censoring, multiple detection limits, variability factor, relative maximum detection limit and unit.

There are a total of 16 censored data sets. Each data set is for a specific urban air toxic emitted from combustion of a specific fuel type. The sample sizes vary from 8 to 29. The censored data of Case Nos. 1, 2, 13 and 14 have a single detection limit, while the others have multiple detection limits. There are nine data sets in which less than 30% of the observations are censored. Four data sets have between 30% and 60% censoring and three have greater than 60% censoring. The latter are defined as highly censored data. The variability factor given in the table is defined as the largest detected value divided by the smallest detected value. The variability factor is an approximate relative indicator of variability in the data. The variability factors range

from 2 to more than 30,000. For Case Nos. 3, 7, 8 and 16, one or more detection limits are larger than the largest detected value.

3.2 Quantification of Inter-Unit Variability and Graphical Evaluation of Goodness-of-fit

The approximation procedure for identifying promising candidate distributions to fit to the censored data was applied to 13 of the 16 emission factor data sets. Data sets for Case Nos. 3, 5 and 8 were not included because the procedure was deemed to be unreliable in cases with a high degree of censoring. Thus, for Case Nos. 3, 5 and 8, all three types of parametric distributions were considered for evaluation in the later step involving graphical analysis. Of the remaining 13 cases, for 12 of them two or more of the three types of parametric distributions fit to the modified data could not be rejected by the K-S test at a 0.05 significance level. For Case 9, for which all of the candidate distributions fit to the modified data were rejected, the difference between the critical and test values of the K-S statistic was considered small for both the lognormal and Weibull distributions; therefore, these two distributions were further evaluated graphically.

The MLE method was used to fit each type of candidate parametric distribution identified based upon the first screening step to each original censored data set. As a second step for evaluating goodness-of-fit, the confidence intervals for the fitted distribution were evaluated using bootstrap simulation, as previously described. The inter-unit variability of the observed data, the detection limits, the fitted parametric distribution, and the confidence intervals of the fitted distribution were compared graphically. The confidence intervals represent uncertainty attributable to random sampling error. Distributions for which the confidence intervals enclose a larger proportion of the data were preferred over those that did not.

The procedure is illustrated with Case 7 as shown in Figures 1(a), 1(b) and 1(c) for the lognormal, gamma and Weibull candidate distributions. The emission factor data contain 29 data points, of which three are censored. Each of the three censored data points has a different detection limit. In order to plot the data as a cumulative distribution function, it is necessary to estimate the rank of each data point. In this case, all of the 26 detected data points are smaller than the largest detection limit. For example, the largest observed data point could have a rank as low as 26 or as high as 29 depending upon the true but unknown values of the non-detected data. Therefore, there is ambiguity regarding what rank to assign to each observed data point, which causes a range of possible ranks and corresponding cumulative probabilities for each of the detected values. The ranges of the cumulative probabilities of the detected points are represented by vertical lines in the figures.

The 50, 90 and 95 percent confidence intervals of the estimated CDF are represented by dark grey, grey and white bands respectively. The CDF fit to the original data is represented by a white dashed line. The 95 percent confidence intervals enclose almost all of the detected points for all three candidate distributions. The exceptions are that portions of the ranges of the possible cumulative probabilities of six detected points for the lognormal distribution, three detected points for the gamma distribution and two detected points for the Weibull distribution are partly outside the 95 percent confidence interval. The lognormal distribution is "tail-heavy" and gives a shorter lower tail but a longer upper tail than the gamma and Weibull distributions. Of the three distributions evaluated, the Weibull distribution had a slightly better fit in that there were fewer data points for which a portion of possible range of cumulative probability extended outside of the 95 percent confidence interval of the CDF. Therefore, Weibull distribution is recommended as the best fit in this case.

For the other cases, a similar comparison of confidence intervals of the CDF to the detected points was applied in order to determine the best fit from the candidate parametric distributions. Generally, all of the candidate parametric distributions describe the data points well. Since the lognormal distribution has a heavier upper tail than the gamma and Weibull distributions, it is typically a better choice for more highly positively skewed data while the Weibull distribution is typically more suitable for less positively skewed data. The recommended distribution type for each case is given in Table 2. To illustrate typical results obtained, several examples are shown graphically. Figure 2 displays results for Case 11 which has a relatively large sample size of 28 and 5 censored data points. Case 9, as shown in Figure 3, has only one detected point larger than the largest detection limit. Case 13 is a single detection limit case as shown in Figure 4. As shown in Figure 5, Case 5 is an example with high censoring degree and for which there is large amount of uncertainty in the non-detected region.

3.3 Estimation of the Mean and Quantification of Its Uncertainty

Based on the MLE/bootstrap simulation method, an asymptotically unbiased best estimate of the mean and the uncertainty in the mean is quantified (Zhao and Frey, 2003). The best estimated mean is the average of the means from the replicates of the CDF in bootstrap simulation. The MLE/bootstrap method was applied to all of the cases in Table 2 for all of their candidate parametric distributions. The 95 percent confidence interval of the mean is shown as a relative percentage from the mean. MLE/bootstrap estimates of the mean and uncertainty in the mean were obtained for all of the candidate distributions for 13 of the emission factor data sets. For three of the emission factor data sets that had a large proportion of censoring, the approximation procedure for evaluation of goodness-of-fit was not applied as previously noted. Thus, for these three cases, all three of the candidate distributions were included for evaluation

using MLE/bootstrap simulation. However, because of numerical instabilities associated with very high proportions of censoring, ranging from 64 to 85 percent for the three empirical data sets, it was not possible to obtain reliable parameter estimates for distributions fit to some of the bootstrap samples. For example, for a Weibull distribution applied to Case 8, 10 out of 500 of the bootstrap parameter estimates were negative. Therefore, as a quality control step, the results of such simulations were not included in the final analysis. Furthermore, for each of these three cases, there was at least one parametric distribution for which reliable results were obtained.

In some cases, all of the candidate distributions were good fits when evaluated graphically, such as for Case Nos. 6, 7, 10, 14, and 16. In such cases, the differences in the best estimated means were within 25% when comparing the results from alternative distributions to that from the preferred distribution. The differences in the absolute upper and lower level of 95 percent confidence interval of the mean were within 10% when comparing the results from alternative distributions to that from the preferred distribution.

For cases in which one of the candidate parametric distributions was clearly a better fit than the others, the mean values and the 95 percent confidence intervals for the mean would differ more substantially, such as for Case Nos. 1, 2, 5, 9, 11, 12, 13, and 15. In most of these cases, the lognormal distribution was a better fit and was associated with a larger mean value and a larger upper bound to the 95 percent confidence interval on the mean than the other candidate distributions.

Thus, the estimates of the mean and of uncertainty in the mean are not strongly sensitive to the choice of parametric distributions when all of the candidates are of approximately comparable goodness of fit. In contrast, these estimates are more strongly sensitive to the choice of distribution when one distribution is clearly better than the others. The estimated mean from

MLE/bootstrap method based upon the lognormal distribution is larger than for the gamma and Weibull distributions and the confidence interval for the mean is also typically wider.

The cases with a 95 percent confidence interval wider than 200 percent of the mean occur for Cases 1, 4, 5, 9, 11 and 12. Cases 1, 9, and 11 have large inter-unit variability. Cases 4, 5 and 12 have relatively small sample sizes. The recommended 95 percent confidence intervals of the means that are narrower than 100 percent occur for Cases 7, 8, 14 and 16. Case 7 has the largest sample size of 29. Cases 8, 14 and 16 have small inter-unit variability.

Case 10 and 12 are similar in terms of sample size, number of detection limits, proportion of censoring, and the estimated variability factor. However, they differ in terms of the relative range of uncertainty for the mean. The data of Case 12 are more positively skewed and are well-described by a lognormal distribution, whereas the data for Case 10 are well-described by a less skewed and less tail-heavy Weibull distribution. Therefore, the 95 percent confidence interval for Case 12 is more than twice as wide, on a relative basis, than for Case 10.

3.4 Comparison of Conventional and MLE/Bootstrap Methods for Mean Estimates

A comparison of the estimated means from the conventional and MLE/Bootstrap methods is shown in Table 3. The variation in mean estimates among the various methods is typically small when the percentage of censoring is small, such as less than about 25 percent. However, for cases in which the largest detection limits are comparable to or larger than the largest observed data, there is more variation in the means estimated from the various methods. For example, both Cases 7 and 9 have 10.3 percent of censoring, but the variation in the means estimated from Case 7 is larger than that from Case 9 since the former has a large detection limit while the latter has small detection limits. In contrast, for cases with a high percentage of censoring, such as Case 8, there can be substantial variation in the mean estimates.

Replacing the non-detects with zero clearly produces estimates of the mean that are biased low, while replacing the non-detects with the detection limit clearly produces estimates of the mean that are biased high. An unbiased estimate of the mean is expected to be enclosed within the range of these two. This range is defined here as a "reference range." The MLE method is asymptotically unbiased as the sample size increases but can be biased high or low for an individual case with a small sample size. Furthermore, the emission factor data usually have a small sample size. The best estimated mean from the MLE/bootstrap method is taken to be a good estimate if its value is enclosed by the "reference range." The best estimated means from MLE/bootstrap method based on the recommended distribution for Case Nos. 3, 4, 6, 7, 8, 10, 15 and 16 are included in the "reference range". For all of the other cases, the 95 percent confidence intervals of the MLE/bootstrap means enclose the "reference range."

3.5 Comparison of Results for Censored Versus Modified Data

In order to gain insight into whether the MLE/Bootstrap method applied to censored data produces estimates of the mean that are substantially different than a potentially simpler approximation method, a comparison was made to results from an alternative approach. In the alternative approach, each non-detect was replaced with one-half of its detection limit, a parametric distribution was fit to describe inter-unit variability, and uncertainty was estimated in the mean using bootstrap simulation. To facilitate comparisons, ratios of results are shown in Table 4 for the estimated means and for the width of the 95 percent confidence intervals for the mean. These results are based upon the preferred distribution for each case as identified in Table 2. Differences in estimates of means between the two approaches of less than 10 percent and differences in the width of confidence intervals of less than approximately 25 percent were deemed not to be substantial. This situation occurs for 10 cases. Conversely, differences in

estimated means of more than 10% were considered substantial, which occurred for Case Nos. 3, 5, 7, 8, 9 and 12. Case Nos. 3, 5 and 8 have greater than 60% censoring. For Case Nos. 3, 7 and 8, the detection limits for some censored values are much larger than the largest detected value. For these three cases, the estimated means for the modified data are much larger than those for the original censored data, which implies a tendency to overestimate the mean for large censoring, large detection limits, or both using the approximation approach. In contrast, for Case Nos. 9, and 12, for which there are multiple detection limits but for which the percentage of censoring is only 23 percent or less and for which the largest detection limit is small in comparison to the largest detected measurement, the means from the approximation approach are underestimated. However, these two cases involve substantial inter-unit variability in the detected values. Thus, some situations are suggested in which the MLE/bootstrap method applied to censored data, rather than modified data, is clearly preferred: (1) large percentage of censoring; and (2) large detection limits relative to observed values. Furthermore, if there is a large proportion of variability in the data and multiple detection limits, the method applied to censored data is likely to be more reliable.

The with of the 95 percent confidence interval for the mean estimated based upon the modified data tends to be narrower than that based upon the original censored data. This implies that failure to accurately account for censoring will typically lead to underestimation of uncertainty in the mean.

There are some exceptions. For example, for Case 8, the width of the 95% confidence interval in the mean for the modified data is much wider than that for the censored data. This special case has a relative maximum detection limit larger than 75, and the uncertainty results are

more sensitive to the large detected data with a fixed value equal to half of the DL than to the large censored data with a possible value ranging from zero to DL.

4.0 Conclusions and Recommendations

This section answers the key questions posed in the introduction. MLE was used to fit parametric distribution to quantify the inter-unit variability in censored urban air toxic emission factors. MLE is asymptotically unbiased and takes into account the presence of non detects. Combined with bootstrap pair simulation, the uncertainty in the statistics and CDF, including the influence of uncertainty associated with censoring itself, can be quantified for censored data.

The simplified conventional methods provide biased estimates of the mean. The results with conventional methods worsen when there are large detection limits comparable to or larger than the largest observed data and for situations involving a large percentage of censoring. The MLE/Bootstrap provided consistent results for censored data with single or multiple detection limits in 16 cases. Even for data with censoring as high as 80 percent, the MLE/Bootstrap method provided reasonable results. For example, the 95 percent confidence interval encloses all the possible cumulative probabilities of the detected data partly or entirely. Furthermore, a point estimate of the mean may inspire a misleading sense of confidence. The MLE/Bootstrap method provides quantitative information about both the range and likelihood of the mean emission factors, which is a basis for probabilistic risk analysis. Although the MLE/bootstrap method is more computationally intensive, because it offers advantages of being asymptotically unbiased and of providing uncertainty estimates, it is recommended for purposes of analyzing censored urban air toxic emission factor data instead of using conventional methods.

The estimates of the mean and uncertainty in the mean are relatively insensitive to the choice of parametric distribution when multiple types of distributions provide comparable

goodness-of-fit. However, as the degree of discrimination with regard to goodness-of-fit becomes more pronounced among alternative distributions, the differences in estimates of the mean and uncertainty in the mean become more substantial. In these latter cases, there is clearly a difference, for example, in results obtained for a lognormal distribution compared to the less tail-heavy gamma or Weibull distributions.

Sample size and variability in the censored data sets influence the estimated uncertainty. The smaller the sample size and the larger the variability in the data, the larger the uncertainty is usually in the mean or in the CDF for the censored data set. The nature of the censoring in a data set also has influence on the estimated mean and on the uncertainty results. In particular, a large percentage of censoring, detection limits large relative to the largest detected value, multiple detection limits, or some combination of these three, typically leads to larger estimates of uncertainty in the mean that could be obtained with simplified approximation procedures. Generally, the uncertainty in the mean is enlarged by censoring compared to otherwise similar data sets that have no censoring.

The width of the recommended 95 percent confidence interval of the mean ranges from 62 percent to 504 percent relative to the mean value for the 16 cases. When the range of uncertainty is large, the confidence intervals are asymmetric because emission factors must be non-negative. The large uncertainty in the censored emission factor data sets suggests that it is important to quantify uncertainty and that this quantified portion of uncertainty should be taken into account when reporting and using censored emission factors.

The best estimated mean from MLE/Bootstrap method maybe biased for a single case when the sample size is small. However, for all of the 16 cases evaluated here, the best estimate means and their uncertainty ranges were consistent with the bounding ranges of results obtained

from conventional methods in which either zero or the detection limit was assigned to each non-detect. Thus, the MLE/bootstrap method is demonstrated to be a viable technique for quantification of inter-unit variability and uncertainty.

5.0 Acknowledgements

The authors acknowledge the support of the Science to Achieve Results (STAR) grants program of the U.S. Environmental Protection Agency, which funded this work under Grant No. R826790. This paper has not been subject to any EPA review and therefore does not necessarily reflect the views of the Agency, and no official endorsement should be inferred.

6.0 References

Bogen, K.R. and Spear, R.C. (1987), "Integrating Uncertainty and Interindividual Variability in Environmental Risk Assessment," *Risk Analysis*, 7(4): 427-436.

Burmaster, D. and Wilson, A. (2000), "Fitting Second-Order Finite Mixture Models to Data with Censored Values Using Maximum Likelihood Estimation," *Risk Analysis*, 20 (2): 261-271.

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

Efron, B. and Tibshirani, R.J. (1993), *An Introduction to Bootstrap*, Monographs on Statistics and Applied Probability, Chapman & Hall: New York.

Elvira, B., Rao, S., and Porter, P.S. (1999), "Identifying Pollution Source Regions using Multiple Censored Data," *Environmental Sciences and Technology*, 33(13): 2273-2277.

EPA (1993a), "Emission Factor Documentation for AP-42 Section 1.1 Bituminous and Subbituminous Coal Combustion," Prepared by Acurex Environmental Corporation, Edward Aul & Associates, Inc and E.H. Pechan and Associates, Inc, for Office of Air Quality Planning and

Standards, Office of Air and Radiation, U.S. Environmental Protection Agency, Research Triangle Park, NC. http://www.epa.gov/ttn/chief/ap42/ch01/bgdocs/b01s01.pdf (accessed by 5/ 2001)

EPA (1993b), "Emission Factor Documentation for AP-42 Section 1.3 Fuel Oil Combustion," Prepared by Acurex Environmental Corporation, Edward Aul & Associates, Inc and E.H. Pechan and Associates, Inc, for Office of Air Quality Planning and Standards, Office of Air and Radiation, U.S. Environmental Protection Agency, Research Triangle Park, NC. http://www.epa.gov/ttn/chief/ap42/ch01/bgdocs/b01s03.pdf (accessed by 5/ 2001)

EPA (2001), "Background Document Report on Revisions to 5th Edition AP-42 Section 1.6 Wood Residue Combustion In Boilers," Prepared by Eastern Research Group, for Office of Air Quality Planning and Standards, Office of Air and Radiation, U.S. Environmental Protection Agency, Research Triangle Park, NC. http://www.epa.gov/ttn/chief/ap42/ch01/bgdocs/b01s04.pdf (accessed by 5/ 2001)

Faraggi, D., Izikson, P. and Reiser, B. (2003), "Confidence intervals for the 50 per cent response dose," *Statistics in Medicine*, 22 (12): 1977-1988.

Frey, H.C. and Bammi, S. (2002), "Quantification of Variability and Uncertainty in Lawn and Garden Equipment NO_x and Total Hydrocarbon Emission Factors," *Journal of the Air & Waste Management Association*, 52(4): 435-449.

Frey, H.C., and Bharvirkar, R. (2002), *Quantification of Variability and Uncertainty. A Case Study of Power Plant Hazardous Air Pollutant Emissions, Chapter - Human and Ecological Risk Analysis,* D. Paustenbach, Ed., John Wiley and Sons: New York.

Frey, H.C. and Rhodes, D.S. (1996), "Characterization and Simulation of Uncertainty Frequency Distributions: Effects of Distribution Choice, Variability, Uncertainty and Parameter Dependence," *Human and Ecological Risk Assessment,* 4(2): 428-468.

Frey, H.C. and Zheng, J. (2002a), "Quantification of Variability and Uncertainty in Utility NO_x Emission Inventories," *Journal of Air and Waste Management Association*, 52(9): 1083-1096.

Frey, H.C. and Zheng, J. (2002b), "Probabilistic Analysis of Driving Cycle-Based Highway Vehicle Emission Factors," *Environmental Science and Technology,* 36(23): 5184-5161.

Gilliom, R.J. and Helsel, D.R. (1986), "Estimation of Distributional Parameters for Censored Trace Level Water Quality Data 1. Estimation Techniques," *Water Resources Research,* 22(2): 135-146.

Haas, C.N. and Scheff, P.A. (1990), "Estimation of Averages in Truncated Samples," *Environmental Sciences and Technology*, 24(6): 912-919.

Lu, H.C. (2003), "Comparisons of Statistical Characteristic of Air Pollutants in Taiwan by Frequency Distribution," *Journal of the Air & Waste Management Association,* 53(5): 608-616.

Morgan, M.G. and Henrion, M. (1990), Uncertainty: A Guide to Dealing with Uncertainty in Qualitative Risk and Policy Analysis, Cambridge University Press: New York, NY.

Newman, M.C., Dixon, P.M., Looney, B.B., and Pinder, J.E. (1989), "Estimating Mean and Variance for Environmental Samples with Below Detection Limit Observations," *Water Resources Bulletin*, 25(4): 905-915.

Ott, W. (1990), "A Physical Explanation of the Lognormality of Pollutant Concentrations," *Journal of the Air and Waste Management Association*, 40: 1378-1383.

Ott, W. (1995), *Environmental Statistics and Data Analysis*, Lewis Publishers, Boca Raton, FL.

Patrick, D. et al. (1994), *Toxic Air Pollution Handbook,* Van Nostrand Reinhold: New York.

Sadiq, R., Husain, T. and Kar, S. (2002), "Chloroform Associated Health Risk Assessment Using Bootstrapping: A Case Study for Limited Drinking Water Samples," *Water Air and Soil Pollution*, 138 (1-4): 123-140.

Seinfeld, J.H. (1986), *Atmospheric Chemistry and Physics of Air Pollution*, John Willey and Sons, New York.

Smith, R.L., French, C.L., Murphy, D.L., and Thompson, R. (1999), "Ranking and Selection of Hazardous Air Pollutants for Listing under Section 112(k) of the Clean Air Act Amendments of 1990," Technical Support Document of Integrated Urban Air Toxics Strategy, office of Air Quality Planning and Standards, Environmental Protection Agency (EPA). http://www.epa.gov/ttn/atw/urban/urbanpg.html (accessed by 10/1/ 2000)

Tilley, L. (2003), Personal communication with Lori Tilley, Regulatory and Environmental Services Department, City of Jacksonville, FL, via email, May 26th, 2003.

Webster, W.J.; Shih, C. (1996), "A Statistically – Derived Metric to Monitor Time-Speed Variability in Practical Emissions Testing," *Proceedings of the Sixth CRC On-Road Vehicle Emissions Workshop*, Coordinating Research Council: Atlanta, GA.

Zhao, Y. and Frey, H.C. (2003) "Quantification of Variability and Uncertainty for Censored Data Sets and Application to Air Toxic Emission Factors," *Risk Analysis*, submitted.

Table 1. Summary of Emission Factor Data for Selected Urban Air Toxics from External Combustion Emission Source Categories

Case No.	Air Toxics	Fuel[a]	N[b]	Percentage of Censoring[c]	Number of DL	Variability Factor[d]	Relative Max. DL[e]	Unit
1	Benzene	C	18	5.6	1	30459	0.07	g/ton
2	Benzene	W	10	10	1	127	0.006	g/ton
3	Benzene	FO	14	78.6	10	5	4.1	10^{-5} g/ liter
4	Formaldehyde	C	14	35.7	5	90	0.09	10^{-2} g/ ton
5	Formaldehyde	FO	14	64.3	8	118	0.27	10^{-3}g/ liter
6	B(a)p	C	8	37.5	3	30	0.74	10^{-5}g/ ton
7	Mercury	C	29	10.3	3	81	3.93	10^{-2}g/ ton
8	Mercury	FO	13	84.6	11	2	75.5	10^{-6}g/ liter
9	Arsenic	C	29	10.3	3	3366	0.004	g/ ton
10	Arsenic	FO	13	23.1	3	48	0.8	10^{-3}g/ liter
11	Cadmium	C	28	17.9	5	984	0.09	10^{-2}g/ ton
12	Cadmium	FO	13	23.1	3	43	0.1	10^{-4}g/ liter
13	Chromium	C	28	3.6	1	1307	0.03	g/ ton
14	Chromium	FO	13	7.7	1	5	0.105	10^{-4}g/ liter
15	Chromium VI	FO	10	50	5	3	0.5	10^{-4}g/ liter
16	Lead	FO	13	30.8	4	11	0.09	10^{-4}g/ liter

[a] Fuel type; C = coal, W = wood waste, FO = fuel oil
[b] Sample size
[c] Percentage of data samples that are censored
[d] Variability Factor of the detected values, represented by the largest detected value divided by the smallest detected value
[e] Relative maximum detection limit, represented by the largest detection limit divided by the largest detected value

Table 2. Results from MLE/Bootstrap Simulation for Candidate Distributions Fit to Urban Air Toxic Emission Factor Data from External Combustion Sources

Case No.	Air Toxics	Fuel	Unit	MLE/Bootstrap		
				Distribution[a]	Average	95% C.I.(%, %)[b]
1	Benzene	Coal	g/ton	**lognormal**	0.95	(-93, 411)
				Weibull	0.45	(-89, 228)
2	Benzene	Wood Waste	g/ton	lognormal	8.60	(-84, 259)
				gamma	4.16	(-74, 87)
				Weibull	4.36	(-72, 98)
3[c]	Benzene	Fuel Oil	10^{-5} g/liter	**gamma**	1.56	(-60, 120)
4	Formaldehyde	Coal	10^{-2} g/ton	**lognormal**	0.74	(-77, 208)
				Weibull	0.70	(-75, 161)
5[c]	Formaldehyde	Fuel Oil	10^{-3}g/liter	gamma	2.11	(-95, 118)
				Weibull	7.05	(-94, 368)
6	B(a)p	Coal	10^{-5}g/ton	**lognormal**	1.40	(-72, 114)
				gamma	1.28	(-70, 91)
				Weibull	1.28	(-70, 94)
7	Mercury	Coal	10^{-2}g/ton	lognormal	3.83	(-28, 39)
				gamma	3.33	(-27, 33)
				Weibull	3.33	(-27, 35)
8[c]	Mercury	Fuel Oil	10^{-6}g/liter	**lognormal**	5.87	(-31, 32)
9	Arsenic	Coal	g/ton	**lognormal**	0.34	(-91, 264)
				Weibull	0.13	(-79, 161)
10	Arsenic	Fuel Oil	10^{-3}g/liter	lognormal	0.16	(-45, 69)
				gamma	0.13	(-47, 57)
				Weibull	0.13	(-47, 59)
11	Cadmium	Coal	10^{-2}g/ton	**lognormal**	2.47	(-62, 156)
				Weibull	1.81	(-55, 105)
12	Cadmium	Fuel Oil	10^{-4}g/liter	**lognormal**	0.56	(-69, 166)
				gamma	0.43	(-64, 99)
				Weibull	0.43	(-64, 104)
13	Chromium	Coal	g/ton	**lognormal**	0.12	(-59, 123)
				gamma	0.10	(-58, 85)
				Weibull	0.10	(-56, 79)
14	Chromium	Fuel Oil	10^{-4}g/liter	**lognormal**	1.03	(-32, 36)
				gamma	0.98	(-33, 31)
				Weibull	0.98	(-33, 31)
15	Chromium VI	Fuel Oil	10^{-4}g/liter	lognormal	0.61	(-70, 235)
				gamma	0.25	(-63, 63)
				Weibull	0.28	(-49, 50)
16	Lead	Fuel Oil	10^{-4}g/liter	gamma	0.17	(-52, 59)
				Weibull	0.17	(-52, 53)

[a] Candidate parametric distributions, the preferred one is shown in **bold**.

[b] 95% confidence interval relative to the mean value.

Table 2. Continued.

[c] For Cases 3, 5 and 8, all the three distributions are candidate parametric distributions. But based on lognormal and Weibull distributions for Case 3, lognormal distribution for Case 5, gamma and Weibull distributions for Case 8, the estimated parameters for some of the bootstrap samples are out of their definition range, thus the results are nor reported.

Table 3. Comparison of Estimated Means from Conventional and MLE/Bootstrap Methods

Case No.	Air Toxics	Fuel[a]	Unit	Means estimated from the conventional methods[b]				Mean from the MLE/Bootstrap Method
				1	2	3	4	
1	Benzene	C	g/ton	0.53	0.53	0.53	0.53	0.95
2	Benzene	W	g/ton	4.53	4.07	4.11	4.11	4.36
3	Benzene	FO	10^{-5} g/ liter	2.52	0.60	3.00	5.50	1.56
4	Formaldehyde	C	10^{-2} g/ ton	1.11	0.70	0.82	0.91	0.74
5	Formaldehyde	FO	10^{-3}g/ liter	5.88	2.10	2.37	2.64	7.05
6	B(a)p	C	10^{-5}g/ ton	1.56	0.99	1.36	1.77	1.40
7	Mercury	C	10^{-2}g/ ton	3.54	3.17	4.20	5.18	3.33
8	Mercury	FO	10^{-6}g/ liter	6.47	1.20	115	229	5.87
9	Arsenic	C	g/ ton	0.17	0.16	0.16	0.16	0.34
10	Arsenic	FO	10^{-3}g/ liter	0.16	0.12	0.13	0.16	0.13
11	Cadmium	C	10^{-2}g/ ton	2.18	1.81	1.85	1.93	2.47
12	Cadmium	FO	10^{-4}g/ liter	0.53	0.41	0.42	0.44	0.56
13	Chromium	C	g/ ton	0.11	0.10	0.10	0.10	0.12
14	Chromium	FO	10^{-4}g/ liter	1.05	0.97	0.98	0.98	1.03
15	Chromium VI	FO	10^{-4}g/ liter	0.46	0.23	0.28	0.31	0.28
16	Lead	FO	10^{-4}g/ liter	0.24	0.17	0.17	0.17	0.17

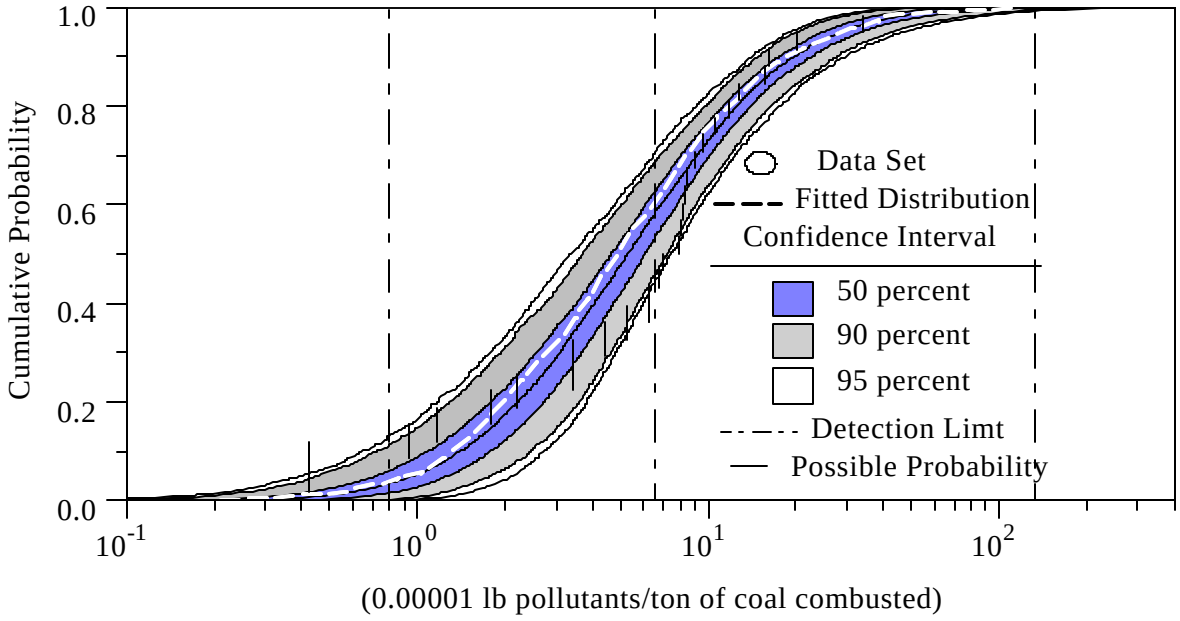
[a] Fuel type; C = coal, W = wood waste, FO = fuel oil

[b] Estimation of mean based upon conventional methods: 1 = removal of non detects; 2 = replace nondetects with zero; 3 = replace nondetects with DL/2; 4 = replace nondetects with DL

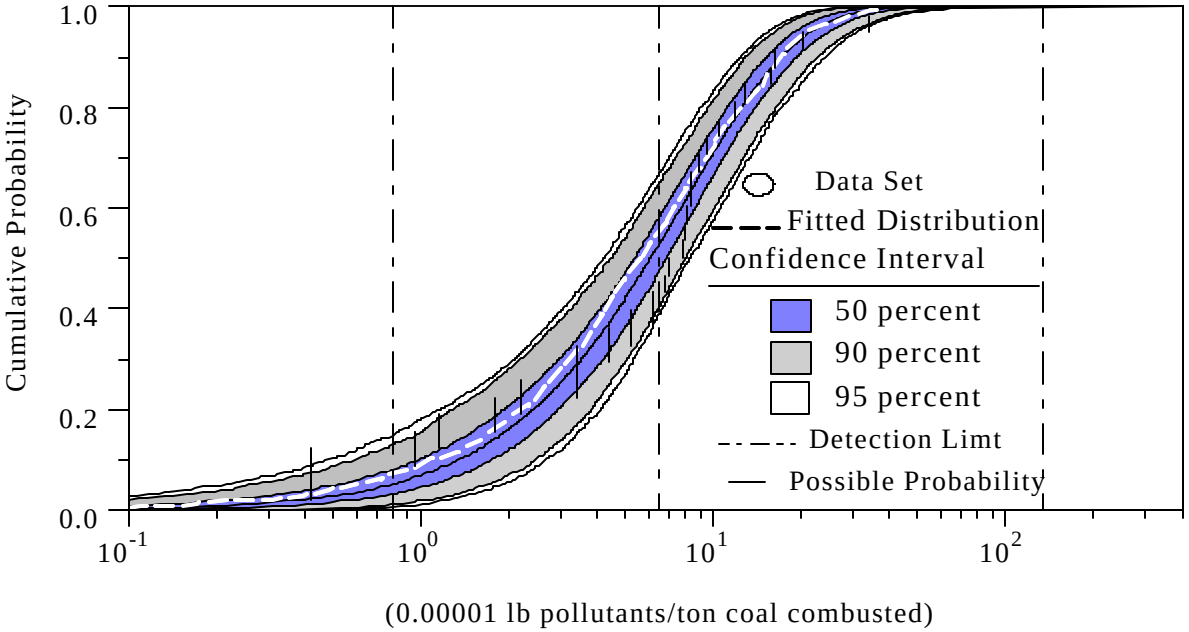
Table 4. Ratio of the Estimated Means and Width of the 95 Perrcent Confidence Intervals for the Means for the MLE/Bootstrap Method Applied to Modified Versus Censored Data.

Case No.	Air Toxics	Fuel	Unit	Ratio of Mean	Ratio of Width of 95% C.I.
1	Benzene	Coal	g/ton	1.09	1.19
2	Benzene	Wood waste	g/ton	0.96	0.94
3	Benzene	Fuel Oil	10^{-5} g/liter	1.92	0.74
4	Formaldehyde	Coal	10^{-2} g/ton	0.94	0.78
5	Formaldehyde	Fuel Oil	10^{-3}g/liter	0.35	0.16
6	B(a)p	Coal	10^{-5}g/ton	1.09	0.88
7	Mercury	Coal	10^{-2}g/ton	1.25	1.69
8	Mercury	Fuel Oil	10^{-6}g/liter	21.0	62.1
9	Arsenic	Coal	g/ton	0.65	0.63
10	Arsenic	Fuel Oil	10^{-3}g/liter	1.00	0.89
11	Cadmium	Coal	10^{-2}g/ton	0.92	0.76
12	Cadmium	Fuel Oil	10^{-4}g/liter	0.89	0.81
13	Chromium	Coal	g/ton	1.00	0.99
14	Chromium	Fuel Oil	10^{-4}g/liter	1.01	1.03
15	Chromium VI	Fuel Oil	10^{-4}g/liter	1.00	0.94
16	Lead	Fuel Oil	10^{-4}g/liter	1.00	0.94

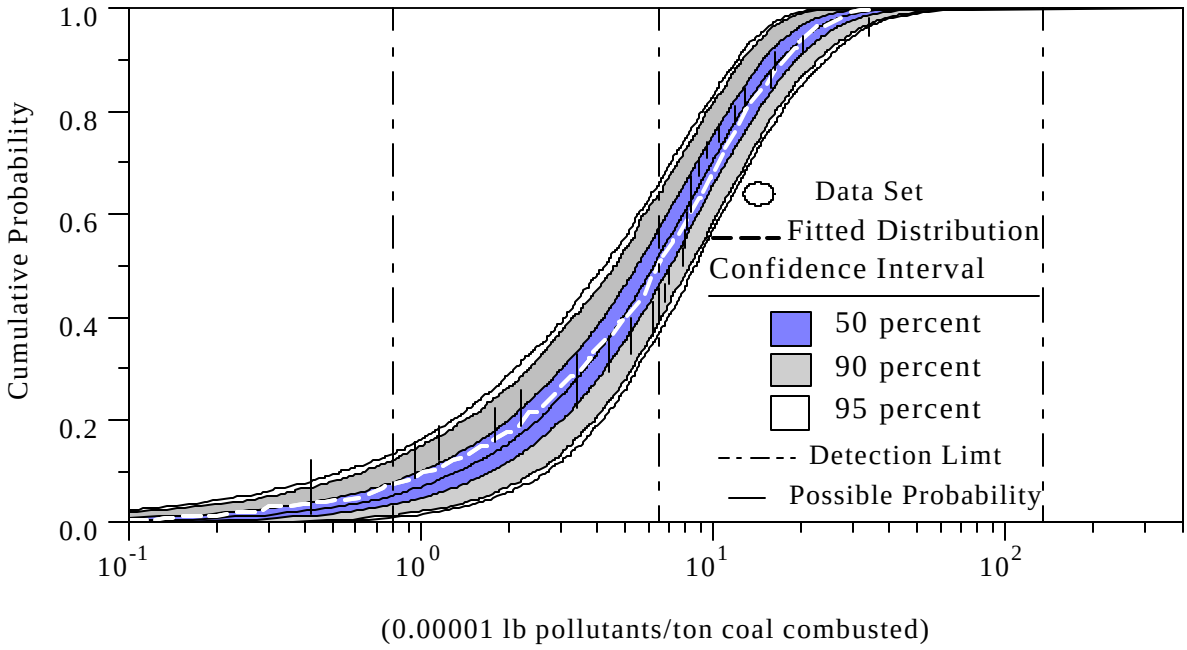
[a.] Ratio of the estimated mean of modified data to that of censored data based on MLE/Bootstrap method



(a) Lognormal Distribution



(b) Gamma Distribution



(c) Weibull Distribution

Figure 1. Variability and Uncertainty in Mercury Emission Factor from Coal Combustion Estimated Based Upon Three Distributions

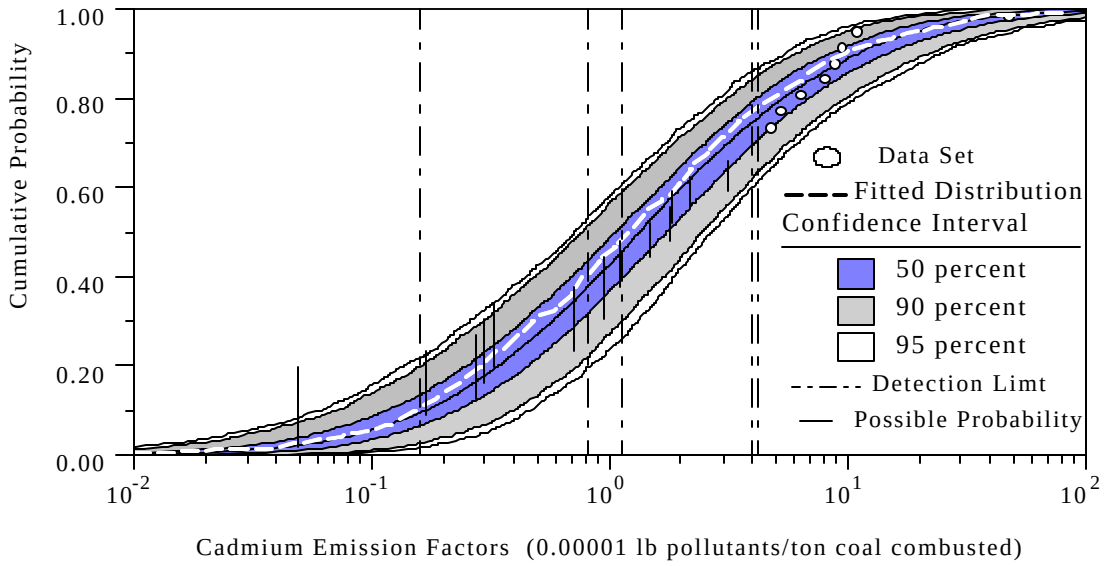


Figure 2. Variability and Uncertainty in Cadmium Emission Factor from Coal Combustion Estimated Based Upon a Lognormal Distribution

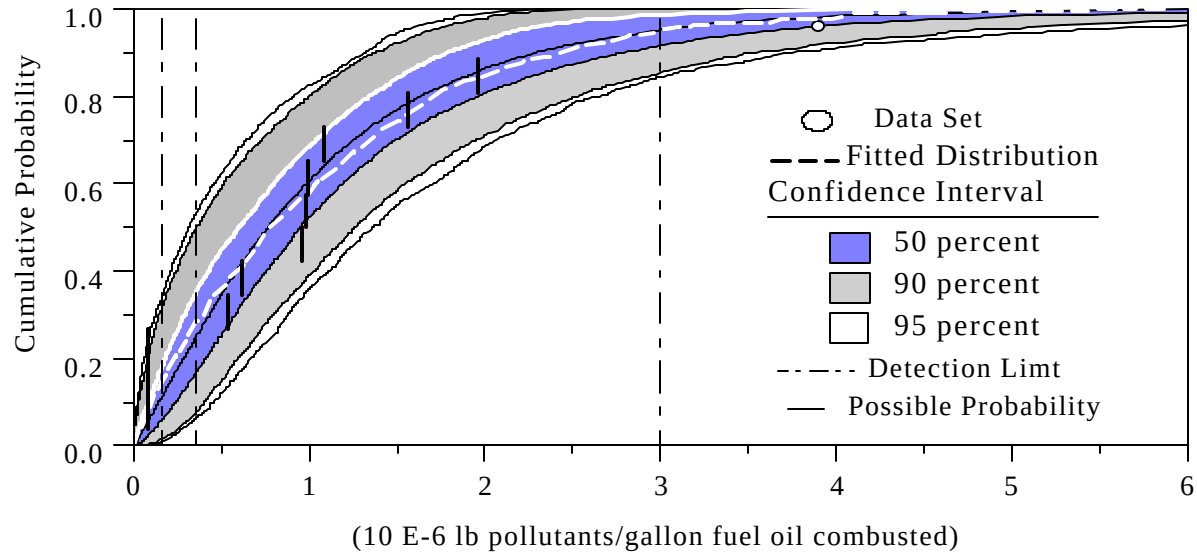


Figure 3. Variability and Uncertainty in Arsenic Emission Factor from Fuel Oil Combustion Estimated Based Upon a Weibull Distribution

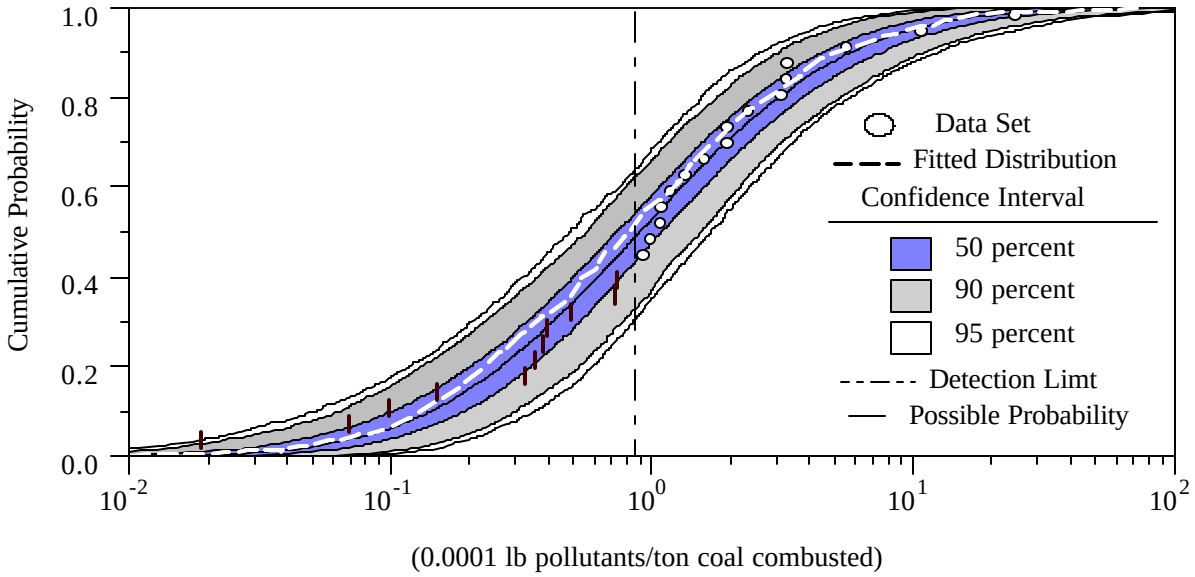


Figure 4. Variability and Uncertainty in Chromium Emission Factor from Coal Combustion Estimated Based Upon a Lognormal Distribution

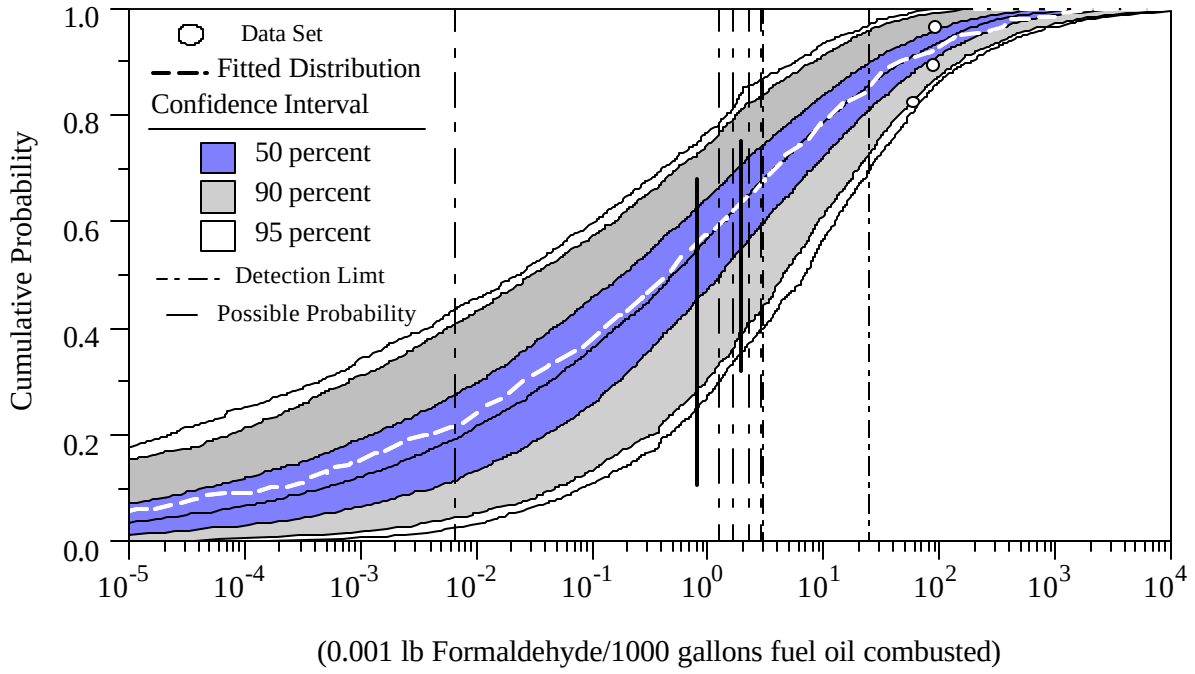


Figure 5. Variability and Uncertainty in Formaldehyde Emission Factor from Fuel Oil Combustion Estimated Based Upon a Weibull Distribution

Part IV

QUANTIFICATION OF VARIABILITY AND UNCERTINTY FOR AIR TOXIC EMISSION INVENTORIES WITH CENSORED EMISSIONS FACTOR DATA

Prepared to be submitted to

Journal of Environment Science and Technology

Abstract. Probabilistic emission inventories were developed for urban air toxic emissions of benzene, formaldehyde, chromium, and arsenic for the example of Houston. Variability and uncertainty in emission factors were quantified for 71 to 97 percent of total emissions, depending upon the pollutant and data availability. Parametric distributions for inter-unit variability were fit using maximum likelihood estimation (MLE) and uncertainty in mean emission factors was estimated using parametric bootstrap simulation. For data sets containing one or more non-detected values, empirical bootstrap simulation was used to randomly sample detection limits for non-detected values and observations for sample values, and parametric distribution for variability were fit using MLE estimators for censored data. Goodness-of-fit for censored data was evaluated using the Kolmogorov-Smirnov test applied to a modified data set and by comparison of cumulative distributions of bootstrap confidence intervals and empirical data. The emission inventory 95 percent uncertainty ranges are as small as minus 25 to plus 42 percent for chromium to minus 75 to plus 224 percent for arsenic with correlated surrogates. Uncertainty was dominated by only a few source categories. Recommendations are made for future improvements to the analysis.

Key Words: Probabilistic emission inventory, Urban air toxics, Emission factors, Censored data, Maximum likelihood estimation, Bootstrap simulation

1.0 Introduction

The purpose of this paper is to demonstrate a method for quantification of variability and uncertainty for cases in which there are a substantial proportion of non-detected data. The methods are illustrated via a case study. The focus of the case study is on selected air toxic pollutants for a specific urban area. The U.S. Environmental Protection Agency has developed a priority list of 33 urban air toxics for additional assessment of the health effects of air toxics in

urban areas.*(1)* An urban area that has been the subject of extensive study by others using deterministic methods is Houston, TX.*(2)* Pollutants for which sufficient data were available to support a probabilistic analysis of air toxic emissions include chromium, formaldehyde, benzene and arsenic, which have risk-related rankings of 5, 6, 10 and 17, respectively, on a nationwide basis, among the listed 33 urban air toxics.

Urban air toxics emissions are subject to both variability and uncertainty. Variability refers to the heterogeneity across different elements of a population over time or space.(3,4) Variability in emissions arises because of differences in feedstocks, ambient conditions, design, or operational practices among facilities.*(5-10)* Uncertainty is lack of knowledge about the true value of a quantity.*(4,11)* Uncertainty in emissions is attributable to random sampling error, measurement error and non-representativeness.*(5-10).* Variability and uncertainty can be quantified simultaneously using a two-dimensional probabilistic framework.*(3,4,6-10,12)*

Emission inventories (EIs) are commonly obtained by the product of emission factors and activity factors. EIs are used by federal, state, and local governments and by private corporations for: (a) characterization of temporal emission trends; (b) emissions budgeting for regulatory and compliance purpose; and (c) prediction of ambient pollutant concentrations using air quality models. If random errors in the EIs are not quantified, erroneous inferences could be regarding trends in emissions, source apportionment, compliance, and the relationship between emissions and ambient air quality.*(13)*

The National Research Council (NRC) recommends that quantifiable uncertainties be addressed in estimating mobile source emission factors *(14)* and logically this recommendation should be extended to other source categories. The NRC has also addressed the need for quantification of uncertainties in emission inventories used in risk assessment.*(15)* Probabilistic

techniques have been applied to estimate uncertainty in emission factors for mobile sources, major stationary sources and area sources, particularly for criteria pollutants (e.g., NO_x) and ozone precursors (e.g., volatile organic compounds).*(5-10,13,16-18)*. Recent work regarding air toxics emissions estimates has focused on situations in which there was only one detection limit.*(10,13)* However, many air toxic data sets have multiple measurements below several different detection limits, since the detection limit is a function of the sample volumes and analytical methods applied separately to each measurement.

The objectives of this paper are: (1) to demonstrate the application of a methodology for quantification of variability and uncertainty in situations involving multiple detection limits; (2) to quantify variability and uncertainty in urban air toxic emission factors for a specific case study; (3) to develop probabilistic EIs for selected pollutants; and (4) to identify key sources in the probabilistic EIs.

2.0 Methodology

Air toxic emission factor data often include one or more measurements below a detection limit. Such data are referred to as censored.*(19)* Conventional methods to dealing with non-detected measurements typically involve replacing non-detected values with zero, one-half of the detection limit, or the detection limit. Such methods lead to biases in estimates of the mean and variance for inter-unit variability in emissions.*(20,21)*. In contrast, the use of maximum likelihood estimation (MLE) to fit parametric probability distributions to non-detected data is asymptotically unbiased.*(20-22)*. The likelihood functions used for candidate parametric distributions are given in the **Supporting Information**.*(22,23)* For censored data, the likelihood function is based upon the cumulative probability of the detection limit, rather than the probability density of an observed value, conditional on parameter estimates. A fitted parametric

distribution is an inference regarding the true but unknown population distribution of inter-unit variability in the emission factor. The uncertainty in the mean or other statistics of the fitted distribution is estimated using bootstrap simulation. Bootstrap simulation is a numerical method for estimation of confidence intervals based upon simulation of random sampling error using Monte Carlo simulation.*(24)*

In order to apply bootstrap simulation to censored data, it is necessary to characterize whether each observation is a detected measurement or below a detection limit. Thus, a binary indicator symbol is used for each observation. For detected values, the binary indicator is set to zero. For non-detected values, the binary indicator is set to one and the corresponding numerical value in the data is the detection limit. Therefore, it is possible to quantify the existence of multiple non-detected values in the data, each of which may have a different detection limit. In bootstrap simulation, the data and indicator symbol pairs are sampled together randomly with replacement n times to generate one bootstrap sample, where n is the sample size of the original data set. The process is repeated B times. To each of the B empirical bootstrap pair samples, the selected type of parametric distribution is fit using MLE. The resulting B realizations of any statistic of interest, such as the mean, variance, or distribution percentiles, characterize the sampling distribution of the statistic associated with random sampling error. These distributions are interpreted as uncertainty in the statistic associated with the variability in the data, sample size, and censoring.*(22)*

2.1 Lognormal, Gamma and Weibull Distributions

For environmental data sets, such as concentrations or emission factors, lognormal, gamma and Weibull distributions are often chosen to represent variability.*(3,6-8,25)* The lognormal distribution is non-negative, positively skewed, and is based upon the Central Limit

Theorem applied to multiplicative processes; therefore, this distribution often well-describes data for physical quantities that arise from multiplicative processes, such as mixing or dilution.(*26,27*). The gamma distribution is non-negative, positively skewed, and similar to the lognormal distribution in many cases but it is less "tail heavy."*(3)* The Weibull distribution is a flexible non-negative distribution that can assume negatively skewed, symmetric, or positively skewed shapes and that has been used to describe air pollutant concentrations.*(3,25)* These three distribution types are used as candidates for describing inter-unit variability in censored air toxics emission factor data.

2.2 Goodness-of-fit Test

Standard methods for evaluation of the goodness-of-fit of a parametric distribution fit to non-censored data include the Kolmogorov-Smirnov (K-S) test.*(3,11,28)* In the case of censored data, a two-step approach to evaluation of goodness-of-fit was used. In the first step, an approximation procedure was used in which each non-detected sample in the data was replaced with one-half of the detection limit to create a modified data set. A parametric distribution was fit to the modified data set using MLE for uncensored data, and the goodness-of-fit was evaluated using the conventional K-S test. If the fitted distribution was not rejected by the K-S test, then that type of parametric distribution model was accepted as a candidate for consideration. In the second step, the selected distribution model was fit to the original censored data and its parameters were estimated using MLE for censored data. Thus, the parameter estimates in the second step are different than those of the first step. The goodness-of-fit in the second step was evaluated by comparing the bootstrap confidence intervals of the fitted cumulative distribution function (CDF) with an empirical distribution of the original sample data. The larger the proportion of data contained within the confidence intervals, the greater the

preference for the particular candidate distribution model. The details of the procedure are illustrated in case studies.

2.3 Monte Carlo Simulation of Uncertainty in the Emission Inventory Model

The emission inventory for a pollutant is given by:

$$\mathrm{EI} = \sum EF_i \times AF_i \qquad (1)$$

Where,

EF_i = emission factor for source i (mass emissions per unit of activity)

AF_i = activity factor for source i (unit of activity)

Based upon selection of the preferred probability distribution model to represent inter-unit variability in the emission factor, uncertainty was estimated for the mean emission factor using bootstrap simulation. Uncertainty in the activity factor was estimated based upon judgment. The uncertainty in the emission inventory was simulated using Monte Carlo simulation *(3,11)*, resulting in an estimate of the probability distribution of uncertainty in the total inventory.

2.4 Identification of Key Sources of Uncertainty

The sensitivity of uncertainty in the total emission inventory for a pollutant to uncertainty in the individual inputs to the inventory was assessed using Spearman correlation coefficients, which measure the strength of the monotonic relationship between two random variables.*(3)* Inputs which had a statistical significant correlation with the outputs were identified as sensitive inputs. The larger the magnitude of the correlation, the greater the sensitivity. Identification of the most highly sensitive inputs enables targeting of resources in future work to collect more or better information in order to reduce uncertainty.

3.0 Houston Emission Inventory

A probabilistic emission inventory for benzene, formaldehyde, chromium and arsenic was developed for Houston based upon the deterministic 1996 inventory. The 1996 inventory was selected because it has been used for a variety of analyses and was the most recent readily available. The focus of the uncertainty analysis was on major source categories. For example, for benzene, the sources emitting more than 20 tons/yr were defined as major source categories. There are 24 major source categories, and these account for 90 percent of the total estimated emissions. For formaldehyde, there are 12 major source categories accounting for 99 percent of total estimated emissions. For chromium and arsenic, there are 27 and 20 major source categories accounting for 71 and 81 percent of the total emissions, respectively. The point estimates of the emissions for each major source category are in Tables 1-4 for benzene, formaldehyde, chromium and arsenic, respectively.

Data for the inter-unit variability in emission factors were identified for each pollutant and source category, where available, based upon information reported by EPA and others, such as the California Air Resources Board and the Coordinating Research Council *(5,7,8,29-39)*. A detailed summary of the sources of data for each source category are given in the **Supporting Information** in Tables S-1 to S-4 for benzene, formaldehyde, chromium and arsenic, respectively. For many source categories, directly relevant data were available via which to estimate inter-unit variability and to infer uncertainty in the mean emission factor. For other source categories, directly relevant data were not available. Therefore, judgments were made regarding surrogates for which uncertainty estimates were likely to be similar. For example, directly relevant sample data were available for inter-vehicle variability in benzene emissions from light duty gasoline vehicles. However, such data were not available for heavy duty

gasoline vehicles. In this case, the relative range of uncertainty in mean benzene emissions for light duty gasoline vehicles was used as a surrogate to estimate the relative range of uncertainty in mean benzene emissions for heavy duty vehicles. A detailed discussion of the judgments made regarding surrogates is given in the Supporting Information.

The emissions for some source categories are estimated as the aggregation of several subcategories. As an example, uncertainty in benzene emissions for petroleum refineries was estimated based upon weights assigned to uncertainties in emissions for gasoline loading racks at bulk terminals and bulk plants, storage losses at a typical gasoline bulk terminal, wastewater treatment, emissions from a typical bulk plant, storage losses at a typical pipeline breakout station, emissions for a typical service station for petroleum refinery, and storage tank for petroleum refinery. For each of the subcategories, data were available from which to quantify relative uncertainty in mean emission rates.

For mobile sources, uncertainty in emissions was estimated based upon the product of uncertainty in the emission rate of total hydrocarbons and of the uncertainty in the percentage of total hydrocarbons emitted as a specific air toxic. Details are shown in the **Supporting Information.**

Directly relevant uncertainty data were available for as little as 45 percent of the major source emissions for formaldehyde to as much as 82 percent in the case of arsenic. When both direct and surrogate uncertainty data are considered, uncertainty was estimated for as little as 71 percent of the nominal emission inventory in the case of chromium to 90 percent or more of the inventories in the cases of benzene and formaldehyde. Thus, for all four pollutants, it was possible to quantify uncertainty for the majority of the emission inventories.

4.0 Results

The variability and uncertainty in the urban air toxics emission factors for different source categories of Houston area were quantified. Probabilistic emission inventories were developed considering the uncertainty in the emission and activity factors. The key sources of uncertainty were identified by sensitivity analysis.

4.1 Quantification of Variability and Uncertainty in Emission Factors

Lognormal, gamma, and Weibull distributions were fit to the available emission factor data for inter-unit variability. For noncensored data, the results of the K-S test were used to choose the best fitting distribution and the adequacy of fit was further assessed graphically based upon comparison of bootstrap confidence intervals of the fitted CDF to the empirical distribution of the data. For censored data, the two step procedure previously described was used to choose the best fitting distribution. In most of the cases, one or more of the distributions could not be rejected by the K-S test and therefore were judged to provide a good fit. There were only 2 cases for benzene, 2 cases for formaldehyde, 1 case for chromium and 1 case for arsenic for which all three of the candidate distribution models were rejected by the K-S test. In those six cases, the model with the smallest K-S test value was selected as the candidate distribution and the fits were judged to be adequate based on graphical comparison of the fitted distribution and its bootstrap confidence intervals with an empirical distribution of the data. The preferred distributions for inter-unit variability are given in Tables 1-4 for benzene, formaldehyde, chromium, and arsenic, respectively, including the parameter estimates. The inter-unit variability in the urban air toxics emission factors is typically large. For example, 13 out of 16 empirical benzene emission factor data sets have a 95 percent probability ratio larger than 2 order-of-magnitude based upon the preferred distribution type as given in the supporting information. The

95% probability ratio is defined as the ratio of the upper level to the lower level of the 95% probability range for inter-unit variability.

Uncertainty in the mean was estimated using parametric bootstrap simulation for the cases with no censoring and with the empirical bootstrap pair approach previously described for cases with censoring. The resulting estimates of the 95 percent confidence intervals for the means are summarized in Tables 1-4.

To illustrate the details of the approach via which variability and uncertainty in an emission factor was quantified; an example case study is given for chromium emission factors for Case 26, industrial residual oil boilers. The emission factor data contain 12 detected values and 1 censored values. All the detected values are larger than the detection limit. The lognormal, gamma and Weibull distributions pass the K-S test and thus are taken as candidate parametric distributions. Simulation of bootstrap pairs coupled with fitting of distributions to each bootstrap pair was used to estimate the confidence intervals of the fitted CDF. Figures 1, 2 and 3 compare an empirical distribution of the data, the CDF for the distribution fitted to the original data, and the bootstrap confidence bands for the cases of a lognormal, a gamma and a Weibull distribution, respectively. The detected points are plotted using the Hazen plotting position and the detection limit is shown for the censored point.

From Figures 1, 2 and 3, both the lognormal, gamma and Weibull distributions are adequate fits. However, the 95 percent confidence interval for the CDF based on the lognormal distribution enclosed the largest proportion of data points compared to the gamma and Weibull distributions. Therefore, the lognormal distribution was selected as the preferred distribution in this case. Thus the resulting 95 percent confidence interval in the emission factor data ranges from minus 32 percent to plus 36 percent relative to the mean value.

As previously described, specific procedures detailed in the supporting information were used for mobile sources and for sources comprised of a weighted combination of subcategories.

4.2 Development of Probabilistic Emission Inventories

Probabilistic emission inventories were developed based upon probabilistic mean emission factors and activity factors. Data regarding uncertainty in the activity were not available. It is expected that there is uncertainty in the activity factors. However, in the absence of empirical data, a judgment was made to assign at least a minimal range of uncertainties to these activity factors. For each source category, a 95 percent confidence interval in the mean of activity factor was assumed ranging from minus 10 percent to plus 10 percent. Therefore, the normalized uncertainty estimates of the activity factors were generated from independent normal distributions with a mean of 1.0 and standard deviation of 0.05. The total uncertainty for each source category was calculated by multiplying the recommended uncertainty estimates of the emission factor by the uncertainty estimates of the corresponding activity factor. The resulting 95 percent confidence interval in the emission inventory of each source category is given in Tables 1 and 4 for benzene, formaldehyde, chromium and arsenic, respectively. Based on the uncertainty of each source category, a probabilistic emission inventory was developed:

$$PEI = \sum_{i=1}^{n} [(UF_{EF,i})(UF_{AF,i})(EI_i)] \quad (2)$$

where,

$UF_{EF, i}$ = Normalized uncertainty factor of emission factors for source i

$UF_{AF, i}$ = Normalized uncertainty factor of activity factors for source i

EI_i = Emission inventory from source i (tons/yr)

In estimating the probabilistic emission inventory, two cases were considered in order to gain insight regarding whether the use of surrogate uncertainty estimates has a significant effect

on the results for uncertainty in the total inventory. In the first case, 100 percent correlation between surrogates was assumed. For example, for benzene the uncertainty in the emission factor for light duty gasoline trucks was correlated with that for light duty gasoline vehicles. In the second case, statistical independence was assumed. The 95 percent uncertainty range for the total inventory for benzene is 2,500 to 9,700 tons/yr versus 2,700 to 8,200 tons/yr when comparing the correlated and uncorrelated surrogates, respectively. For formaldehyde, chromium, and arsenic, there was no difference in the ranges between the two cases. For formaldehyde, the range of uncertainty in either case was from 1,700 to 4,600 tons. For chromium, the range was 3.7 to 7.0 tons/yr and for arsenic the range was 0.6 to 7.1 tons/yr. Thus, there was not a substantial difference when comparing correlated versus uncorrelated surrogates, which indicates that the source categories for which surrogate data were used are not the most important contributors to overall uncertainty.

4.3 Sensitivity Study to Identify the Key Sources of Uncertainty

The rank correlation between the uncertainties in total emissions and the uncertainty in the emission factors of each source category was calculated considering both correlated and uncorrelated surrogates. The results are given in Tables 1-4 for the four pollutants.

For benzene, gasoline onroad mobile sources are the dominate source of uncertainty in the inventory. For formaldehyde, the onroad and nonroad mobile sources are the key sources of uncertainty. For chromium, the key sources of uncertainty are chemical manufacturing-fuel fired equipment-process heaters, external utility coal combustion boilers and hard chromium electroplating. For arsenic, external coal combustion utility boilers are the dominate source of uncertainty in the inventory. The other statistically significant correlations are smaller than 0.2 for each of the pollutants, indicating only weak sensitivity.

5.0 Discussion

Probabilistic emission inventories based primarily on uncertainty in emission factors were developed for benzene, formaldehyde, 1,3-butadiene, and chromium using the 1996 Houston inventory as a basis. A methodology for quantification of uncertainty in the mean emission factors was demonstrated for situations involving censored data. For many source categories, directly relevant data were not available and judgments were made regarding surrogate relative uncertainty estimates. For some source categories, it was necessary to weight data from subcategories; however, the results for uncertainty in the total inventory were not sensitive to judgments regarding these weights. In particular, either the ranges of uncertainty were similar among many of the subcategories or the source category was not important with regard to overall uncertainty in the inventory.

The key characteristics of the probabilistic analysis include: (1) large ranges of inter-unit variability in emission factors for specific source categories; (2) mean emission factor uncertainty ranges from as small as approximately plus or minus 10 percent to as large as -99 to plus 600 percent; (3) relative uncertainty ranges in total emissions ranging from as small as approximately minus 20 to plus 34 percent, as in the case of chromium, to as large as minus 69 to plus 203 percent, in the case of arsenic; and (4) identification of a small number of key sources of uncertainty for each pollutant. Better data collection and reporting should be prioritized for the key source categories.

The quantified ranges of uncertainty for benzene, formaldehyde, chromium and arsenic emissions in the Houston area take into account random sampling error and measurement error in emission factors. The former is influenced by the sample size and inter-unit variability for each emission source category. Random measurement error is accounted for because the observed

variability in the data includes both the true variability and the random component of measurement error, which in turn influences the range of the sampling distribution of the mean.*(40)*

The averaging times of the emissions measurements vary among the source categories and in many cases are not documented in the references from which the data were obtained. It is likely that most of the measurements are for relatively short averaging times on the order of minutes (e.g., for some mobile sources) to perhaps days (e.g., stack testing). Although the desired averaging time for exposure assessment purposes is one year or longer, the uncertainty in the mean emissions estimates is influenced by the limited averaging time of the available data.

The probabilistic emission inventory developed here could be improved in several ways pending availability of additional data or the incorporation of a more extensive expert elicitation component. For example, although biases in the mean emission factors are suspected, especially for fugitive emissions and as a result of process upset, insufficient data were available via which to quantify such biases. Other possible sources of bias include lack of representative data (e.g., measurements may have been for load or operating conditions not typical of annual average in-use activity) and the use of surrogate data for source categories in which data were lacking or not readily available. Expert elicitation could be used to encode judgments regarding the additional uncertainty associated with nonrepresentative or surrogate data. As new data become available, the assessment can be updated. A key obstacle to quantification of uncertainty based upon statistical data analysis is obtaining the necessary data. Often, data are measured and reported by multiple organizations. In the long term, the development of a protocol for archiving such data and making the data available would facilitate probabilistic analysis.

The uncertainty in the activity factors here is based on an approximate judgment, mostly as an acknowledgement that uncertainty exists and as a placeholder pending better information. In the long term, the quantifiable uncertainty in the activity factors should be incorporated based upon expert judgment.

The results of this work demonstrate that random sampling error and measurement error in emission factors are substantial sources of quantifiable uncertainty in the emission inventories of benzene and formaldehyde in the Houston area. The positively skewed ranges of uncertainty appropriately account for the fact that emissions must be non-negative. The MLE/bootstrap methodology used here provides asymptotically unbiased estimates of the mean, including for cases that involve non-detected data. The substantial ranges of uncertainty estimated here should be taken into account when conducting air quality modeling and exposure assessment. Furthermore, the identification of key sources of uncertainty in the inventory serves as an aid to prioritizing resources for additional data collection or research in order to reduce uncertainty.

6.0 Acknowledgements

The authors acknowledge the support of the U.S. Environmental Protection Agency, which funded portions of this work via MCNC (now the Carolina Environmental Program). The statistical methods used in this report were developed under U.S. EPA STAR Grant Number R826790. Jawad Touma of EPA, Adel Hanna, Jeff Vukovich, and Sarav Arunachalam of the Carolina Environmental Program, and Steve Hanna of Hanna Consultants provided useful suggestions and information. Although the research described in this paper has been funded wholly or in part by the U.S. EPA, this paper has not been subject to any EPA review and therefore does not necessarily reflect the views of the Agency, and no official endorsement should be inferred.

7.0 Supporting Information

The supporting information (SI) contains text, tables, and figures pertaining to details of the MLE parameter estimation method for censored data, specific sources of data for each emission factor, the methods used for quantification of uncertainty in mobile source emission factors and for source categories based upon aggregation of subcategories, and a graphical summary of the relative range of uncertainty in total emissions for each pollutant for both correlated and uncorrelated surrogates.

8.0 References

1. Smith, R.L., French, C.L., Murphy, D.L., and Thompson, R, *Ranking and Selection of Hazardous Air Pollutants for Listing under Section 112(k) of the Clean Air Act Amendments of 1990*, Technical Support Document of Integrated Urban Air Toxics Strategy, by EPA office of Air Quality Planning and Standards, 1999. http://www.epa.gov/ttn/atw/urban/urbanpg.html (accessed by 10/1/ 2000)

2. "Example Application of Modeling of Toxic Air Pollutants in Urban Areas," Office of Air Quality Planning and Standards, U.S. Environmental Protection agency, Research Triangle Park, 2002. Report No. EPA-454/R-02-003.

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

4. Bogen, K.R. and Spear, R.C., "Integrating Uncertainty and Interindividual Variability in Environmental Risk Assessment," *Risk Analysis*, 7(4): 427-436.

5. Frey, H.C. and Zheng, J. (2002), "Probabilistic Analysis of Driving Cycle-Based Highway Vehicle Emission Factors," *Environmental Science and Technology,* 36(23): 5184-5161.

6. Frey, H.C. and Zheng, J. (2002), "Quantification of Variability and Uncertainty in Utility NO_x Emission Inventories," *Journal of Air and Waste Management Association,* 52(9): 1083-1096.

7. Frey, H.C. and Bammi, S. (2002), "Quantification of Variability and Uncertainty in Lawn and Garden Equipment NO_x and Total Hydrocarbon Emission Factors," *Journal of the Air & Waste Management Association,* 52(4): 435-449.

8. Frey, H.C. and Bammi, S. (2003), "Probabilistic Nonroad Mobile Source Emission Factors," *Journal of Environmental Engineering,* 129(2): 162-168.

9. Frey, H.C.; Rhodes, D.S., *Human Health and Ecological Risk Assessment*, **1996**, 2:762-797.

10. Frey, H. C., and Bharvirkar, R. Chapter in *Risk Assessment of Environmental and Human Health Hazards: A Textbook of Case Studies,* D. Paustenbach, Ed., John Wiley and Sons: New York, 2001.

11. Morgan, M.G. and Henrion, M. (1990) *Uncertainty: A Guide to Dealing with Uncertainty in Qualitative Risk and Policy Analysis,* Cambridge University Press: New York, NY.

12. Frey, H.C. and Rhodes, D.S. (1998), "Characterization and Simulation of Uncertainty Frequency Distributions: Effects of Distribution Choice, Variability, Uncertainty and Parameter Dependence," *Human and Ecological Risk Assessment,* 4(2): 428-468.

13. Frey, H.C., Bharvirkar, R. and Zheng J., *Quantitative Analysis of Variability and Uncertainty in Emissions Estimation,* Prepared by NC State University for Office of Air Quality Planning and Standards, U.S. Environmental Protection agency, Research Triangle Park, NC, 1999.

14. NRC, *Modeling Mobile Source Emissions,* National Academy Press: Washington D.C., 2000.

15. NRC, *Science and Judgment in Risk Assessment,* National Academy Press: Washington D.C., 1994.

16. McCleese, D.L., and LaPuma, P.T. (2002), "Using Monte Carlo Simulation in Life Cycle Assessment for Electric and Internal Combustion Vehicles," *International Journal of Life Cycle Assessment*, 7(4): 230-236.

17. Averill, A.F., Ingram, J.M., and Nolan, P.F. (1999), "A Study of the Dispersion of Solvent Vapour in the Workspace During Wipe Cleaning of Metal Components with Organic Solvents – A Monte Carlo Uncertainty Analysis," *Transaction of the Institute of Metal Finishing*, 77(9): 204-208.

18. Pollack, A.K.; Bhave P.; Heiken J.; Lee K.; Shepard S.; Tran C.; Yarwood G.; Sawyer R.F.; Joy B.A. *Investigation of Emission Factors in the California EMFAC7G Model*; PB99-149718INZ; Coordinating Research Council: Atlanta, GA, 1999.

19. Rao, S.T., Ku, J.Y., and Rao, K. S. (1991), "Analysis of Toxic Air Contaminant Data Containing Concentrations Below the Limit of Detection," *Journal of the Air and Waste Management Association,* 41(4): 442-448.

20. Haas, C.N. and Scheff, P.A. (1990), "Estimation of Averages in Truncated Samples," *Environmental Sciences and Technology*, 24(6): 912-919.

21. Newman, M.C., Dixon, P.M., Looney, B.B., and Pinder, J.E. (1989), "Estimating Mean and Variance for Environmental Samples with Below Detection Limit Observations," *Water Resources Bulletin*, 25(4): 905-915.

22. Zhao, Y., and H.C. Frey (2003) "Quantification of Variability and Uncertainty for Censored Data Sets and Application to Air Toxic Emission Factors," *Risk Analysis*, submitted and undergoing revisions in response to reviewer comments.

23. Cohen, A.C. and Whitten, B. (1988), *Parameter Estimation in Reliability and Life Span Models*, M. Dekker: New York.

24. Efron, B. and Tibshirani, R.J., An Introduction to Bootstrap, Monographs on Statistics and Applied Probability, Chapman & Hall: New York, 1993.

25. Seinfeld, J.H., *Atmospheric Chemistry and Physics of Air Pollution*, John Willey and Sons, New York, 1986.

26. Ott, W. (1990), "A Physical Explanation of the Lognormality of Pollutant Concentrations," *Journal of the Air and Waste Management Association*, 40: 1378-1383.

27. Ott, W. (1995), *Environmental Statistics and Data Analysis*, Lewis Publishers, Boca Raton, FL.

28. Lu, HC., "Comparisons of Statistical Characteristic of Air Pollutants in Taiwan by Frequency Distribution," *Journal of the Air & Waste Management Association*, 53 (5): 608-616.

29. EPA, *Locating and Estimating Air Emissions from Sources of Formaldehyde*, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1991. http://www.epa.gov/ttn/chief/le/benzene/index.html,

30. EPA, *Emission Factor Documentation for AP-42 Section 1.1 Bituminous and Subbituminous Coal Combustion*, by Acurex Environmental Corporation, Edward Aul & Associates, Inc and E.H. Pechan and Associates, Inc, for Office of Air Quality Planning and Standards, Office of Air and Radiation, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1993. http://www.epa.gov/ttn/chief/ap42/ch01/bgdocs/b01s01.pdf

31. EPA, *Emission Factor Documentation for AP-42 Section 1.3 Fuel Oil Combustion*, by Acurex Environmental Corporation, Edward Aul & Associates, Inc and E.H. Pechan and Associates, Inc, for Office of Air Quality Planning and Standards, Office of Air and Radiation, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1993. http://www.epa.gov/ttn/chief/ap42/ch01/bgdocs/b01s03.pdf

32. EPA, *AP-42, Fifth Edition, Volume 1, Chapter 1: External Combustion Sources, 1.6 Wood Residue Combustion in Boilers*, 1998. http://www.epa.gov/ttn/chief/ap42/ch01/

33. EPA, *Locating and Estimating Air Emissions from Sources of Benzene*, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1998. http://www.epa.gov/ttn/chief/le/benzene/index.html.

34. CARB, *California Air Toxics Emission Factor II Database*, Research Division, Air Resources Board, California Environmental Protection Agency, 2000. http://www.arb.ca.gov/emisinv/catef/catef.htm.

35. CRC, *Heavy-duty Gasoline Vehicle Chassis Dynamometer Testing for Emissions Inventory*, Coordinating Research Council, 2003. http://www.crcao.com/.

36. IPCC, *Background Papers: IPCC Expert Meetings on Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories*, by Intergovernmental Panel on Climate Change, National Greenhouse Gas Inventories Programme, Institute for Global Environmental Strategies, 2108-11, Kamiyamaguchi, Hayama, Kanagawa, Japan, 2002.

37. Bammi, S., *Quantitative Analysis of Variability and Uncertainty in On-Road and Non-Road Mobile Source Emission Factors*, Master thesis, Department of Civil Engineering, North Carolina State University, Raleigh, NC, 2001. www.lib.ncsu.edu

38. EPA, *Emission Factor Documentation for AP-42 Section 2.6 Medical Waste Incineration*, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1993. http://www.epa.gov/ttn/chief/ap42/ch02/bgdocs/b02s03.pdf

39. EPA, "Locating and Estimating Air Emissions from Sources of Chromium," Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1984. http://www.epa.gov/ttn/chief/le/chromium.pdf.

40. Zheng, J. (2002), "Quantification of Variability and Uncertainty in Emission Estimation: General Methodology and Software Implementation," Ph.D. Dissertation, Department of Civil Engineering, North Carolina State University, Raleigh. www.lib.ncsu.edu

Table 1. Quantification of Variability and Uncertainty for Benzene Emission Inventory

Case No.	Emission Source Description	EI[a] (tons/yr)	n[b]	Variability in Emission Factor[c]	Uncertainty in Emission Factor (%, %)[d]	Uncertainty in Emission Inventory (%, %)[e]	Rank Correlation 1[f]	Rank Correlation 2[g]
1	Mobile Source – Light Duty Gasoline Vehicles	1164.40	---	---	(-87, 236)	(-88, 249)	**0.909**	**0.672**
2	Mobile Source – Light Duty Gasoline Trucks	846.19	---	---	(-87, 236)	(-88, 252)	**0.909**	**0.514**
3a	Gasoline loading racks at bulk terminals and bulk plants	---	3	L (-3.62, 1.40)	(-89, 345)	(-90, 327)	---	---
3b	Storage losses at a typical gasoline bulk terminal (non Winter)	---	11	L (-3.86, 1.44)	(-72, 185)	(-72, 212)	---	---
3c	Storage losses at a typical gasoline bulk terminal (winter)	---	11	L (-3.53, 1.43)	(-71, 195)	(-74, 196)	---	---
3d	For a typical bulk plant	---	6	L (-3.02, 1.67)	(-86, 342)	(-87, 349)	---	---
3e	Storage losses at a typical pipeline breakout station (non-winter)	---	11	L (-2.97, 1.77)	(-84, 226)	(-84, 284)	---	---
3f	Storage losses at a typical pipeline breakout station (winter)	---	11	L (-2.64, 1.75)	(-82, 285)	(-84, 279)	---	---
3g	For typical Service Station for petroleum refinery	---	7	L (-3.93, 1.35)	(-78, 213)	(-77, 213)	---	---
3h	Petroleum Refinery wastewater	---	19	G (0.53, 65.54)	(-52, 71)	(-53, 70)	---	---
3i	Storage tank for petroleum refinery	---	5	L (4.48, 2.57)	(-98, 562)	(-99, 596)	---	---
3	Petroleum refinery	714.30	---	---	(-55, 158)	(-55, 170)	**0.169**	**0.215**
4	4-stroke lawn and garden engines	686.87	---	---	(-34, 46)	(-33, 50)	**0.125**	**0.137**
5	2-stroke lawn and garden engines	234.01	---	---	(-32, 40)	(-32, 41)	0.057	0.055
6	Construction, farm and industrial engine (diesel 4 Stroke)	142.48	---	---	(-26, 30)	(-28, 32)	0.023	0.023
7	Oil and natural gas production	140.62	---	---	(-10, 10)	(-13, 14)	0.066	0.053

(Continued)

Table 1. Continued

Case No.	Emission Source Description	EI[a] (tons/yr)	n[b]	Variability in Emission Factor[c]	Uncertainty in Emission Factor (%, %)[d]	Uncertainty in Emission Inventory (%, %)[e]	Rank Correlation 1[f]	Rank Correlation 2[g]
8	Storage and transport, Natural Gas Transmissions and Marine Transport	100.38	---	---	(-10, 10)	(-14, 14)	0.066	-0.007
9	Mobile source-Heavy Duty Gasoline Vehicle	78.95			(-87, 236)	(-88, 255)	**0.909**	0.057
10	Other combustion-forest wildfires	54.40	6	W (1.82, 1.61)[h]	(-40, 45)	(-47, 50)	0.044	0.031
11	Solid waste disposal- sewage treatment	49.47	16	L (0.32, 3.27)	(-98, 328)	(-98, 323)	0.015	**0.098**
12	Industrial Processes; Chemical Manufacturing; Acetylene production	47.85	---	---	(-10, 10)	(-14, 13)	0.066	-0.021
13	Fuel oil external combustion	45.32	14 (11)	G (0.28, 3.27)	(-68, 120)	(-68, 122)	0.051	0.021
14	Typical ethylene plant	43.29	8	L (-4.18, 2.85)	(-99, 221)	(-99, 216)	0.015	0.026
15	Storage and Transport; Petroleum Product Storage; Gasoline Service Stations; Stage 1: Total	40.09	---	---	(-78, 213)	(-73, 180)	0.043	**0.121**
16	Industrial Processes; Petroleum Industry; Fugitive	38.64	---	---	(-10, 10)	(-13, 14)	0.066	0.019
17	Other combustion-managed prescribed burning	33.92	7	G (2.21, 0.59)[h,i]	(-43, 53)	(-45, 52)	0.038	0.076
18-1	Total Hydrocarbon from Heavy Duty Diesel Vehicle	---	24	L (0.46, 1.46)	(-58, 150)	---		
18-2	Benzene fraction in THC from Heavy Duty Diesel Vehicle	---	24	W (3.36, 0.009)	(-13, 13)	---		

(Continued)

Table 1. Continued

Case No.	Emission Source Description	EI[a] (tons/yr)	n[b]	Variability in Emission Factor[c]	Uncertainty in Emission Factor (%, %)[d]	Uncertainty in Emission Inventory (%, %)[e]	Rank Correlation 1[f]	Rank Correlation 2[g]
18	Benzene emission factor from Heavy Duty Diesel Vehicle	33.87	---	---	(-59, 166)	(-59, 182)	-0.028	0.021
19	Industrial Processes; Chemical Manufacturing; Fugitive Emissions	30.47	---	---	(-10, 10)	(-14, 14)	0.066	-0.026
20	Mobile source-aircraft	26.15	---	---	(-51, 72)	(-50, 79)	**0.099**	**0.086**
21	Industrial Processes; Petrolem Industry; Fugitive Emissions; Miscellaneous: Sampling/Non-Asphalt	26.00	---	---	(-10, 10)	(-13, 13)	0.066	-0.003
22	Petroleum refinery-process vent in refinery product	24.89	13	G (0.32, 2.37)	(-73, 113)	(-72, 127)	-0.065	-0.007
23	Loading, ballasting and transit losses from marine vessels	21.59	9	L (-4.17, 0.54)	(-31, 37)	(-33, 39)	**0.105**	**0.107**
24	Industrial Processes; Chemical Manufacturing; Processes; Fugitive leaks	20.73	---	---	(-10, 10)	(-13, 14)	0.066	0.077

a. Point estimate of benzene emission inventory

b. Sample size, for censored data set, the number of non-detects is shown in parenthesis

c. Inter-unit Variability in emission factor. L: lognormal distribution; G: gamma distribution; W: Weibull distribution. The parameters of the distribution are given in parenthesis

d. The 95 % confidence interval relative to the mean is given.

e. The 95% confidence interval relative to the mean is given. For the cases in which there is no information of variability in emission factors while there is information of uncertainty in the emission factor, the latter is based upon surrogate or previous work.

f. Rank correlation between the uncertainty in the total emissions and the uncertainty in the emission factors for each source category with correlated surrogates. Statistically significant correlations are shown in boldface.

Table 1. Continued

[g] Rank correlation between the uncertainty in the total emissions and the uncertainty in the emission factors for each source category with uncorrelated surrogates. Statistically significant correlations are shown in boldface.

[h] Rejected by the K-S test, but the fit is judged to be adequate.

[i] Fit parametric distribution with MOMM method instead of MLE since MOMM results in a smaller K-S test

Table 2. Quantification of Variability and Uncertainty for Formaldehyde Emission Inventory

Case No.	Emission Source Description	EI[a] (tons/yr)	n[b]	Variability in Emission Factor[c]	Uncertainty in Emission Factor (%, %)[d]	Uncertainty in Emission Inventory (%, %)[e]	Rank Correlation 1[f]	Rank Correlation 2[g]
1a	Nonroad - 4-stroke lawn and garden engines	182.79	---	---	(-39, 59)	(-40, 56)	---	---
1b	Nonroad - 2-stroke lawn and garden engines		---	---	(-36, 51)	(-37, 53)	---	---
1c	Nonroad-CFI engine (diesel 4S)	934.99	---	---	(-32, 43)	(-35, 44)	---	---
1d	Nonroad-Aircraft	163.75	---	---	(-53, 80)	(-53, 83)	---	---
1	Noroad mobile source	1281.54	---	---	(-26, 35)	(-29, 35)	**0.306**	**0.305**
2a	Onroad gasoline	763.53	---	---	(-87, 224)	(-88, 242)	---	---
2b-1	Onroad diesel engines-THC	---	24	L (0.46, 1.46)	(-58, 150)		---	---
2b-2	Onroad diesel engines-formaldehyde fraction	---	24	W (1.43, 0.086)[h]	(-28, 30)		---	---
2b	Onroad diesel engines	212.55	---	---	(-63, 166)	(-62, 157)	---	---
2	Onroad mobile source	976.08	---	---	(-75, 177)	(-76, 176)	**0.892**	**0.893**
3a	Stationary reciprocating internal combustion engines (material type: liquid)	---	12	L (-2.21, 1.58)	(-74, 217)	(-76, 222)	---	---
3b	Stationary reciprocating internal combustion engines (material type: gas)	---	12	L (1.12, 2.59)	(-96, 409)	(-96, 408)	---	---
3	Internal combustion engines	143.92	---	---	(-77, 269)	(-76, 179)	**0.136**	**0.136**
4	Oil and gas extraction	99.51	---	---	(-10, 10)	(-13, 13)	0.030	0.024
5	Chemical and allied processes	69.71	---	---	(-10, 10)	(-13, 14)	0.030	0.006
6a	Stationary combustion turbines (material type: liquid)	---	3	L (-3.18, 0.38)	(-36, 50)	(-36, 47)	---	---
6b	Stationary combustion turbines (material type: gas)	---	10	L (-0.12, 0.98)[h]	(-55, 100)	(-53, 98)	---	---

(Continued)

Table 2. Continued

Case No.	Emission Source Description	EI[a] (tons/yr)	n[b]	Variability in Emission Factor[c]	Uncertainty in Emission Factor (%, %)[d]	Uncertainty in Emission Inventory (%, %)[e]	Rank Correlation 1[f]	Rank Correlation 2[g]
6	Combustion turbines	66.62			(-36, 56)	(-37, 58)	0.072	0.071
7	Petroleum refineries	64.09	3	L (0.057, 0.58)	(-51, 93)	(-52, 80)	**0.099**	**0.097**
8	Open burning, forest and wildfires	39.92	---	---	---	---	---	---
9	Open burning, prescribed burnings	24.74	---	---	---	---	---	---
10	Utility boilers	7.59	---	---	(-55, 152)	(-62, 154)	**0.086**	0.083
11a	External combustion- coal combustion	---	14 (5)	L (-0.59, 1.44)	(-77, 208)	(-77, 209)	---	---
11b	External combustion-wood fired waste	---	20	L (1.11, 1.34)	(-58, 129)	(-57, 119)	---	---
11c	External combustion-fuel oil	---	14 (9)	W (0.24, 1.42)	(-94, 368)	(-93, 367)	---	---
11	Industrial boilers	5.06	---	---	(-55, 152)	(-60, 151)	**0.086**	0.041
12	Structure fires	4.78	---	---	---	---	---	---

a. Point estimate of formaldehyde emission inventory

b. Sample size, for censored data set, the number of non-detects is shown in parenthesis

c. Variability in emission factor. L: lognormal distribution; G: gamma distribution; W: Weibull distribution. The parameters of the distribution are given in parenthesis

d. The 95 % confidence interval relative to the mean is given.

e. The 95% confidence interval relative to the mean is given. For the cases in which there is no information of variability in emission factors while there is information of uncertainty in the emission factor, the latter is based upon surrogate or previous work.

f. Rank correlation between the uncertainty in the total emissions and the uncertainty in the emission factors for each source category with correlated surrogates. Statistically significant correlations are shown in boldface

g. Rank correlation between the uncertainty in the total emissions and the uncertainty in the emission factors for each source category with uncorrelated surrogates. Statistically significant correlations are shown in boldface.

h. Rejected by K-S test, but the fit is judged to be adequate

Table 3. Quantification of Variability and Uncertainty for Chromium Emission Inventory

Case No.	Emission Source Description	EI[a] (tons/yr)	n[b]	Variability in Emission Factor[c]	Uncertainty in Emission Factor (%, %)[d]	Uncertainty in Emission Inventory (%, %)[e]	Rank Correlation 1[f]	Rank Correlation 2[g]
1	PETROLEUM REFINERIES - CATALYTIC CRACKI	1.868	3	W (121.8, 5.01)	(-1.5, 1.0)	(-10, 9.7)	0.092	0.090
2	Marine Vessels, Commercial	1.026						
3	CHEM MFG-FUEL FIRED EQ-PROCESS HEATERS	0.810	3	L (-2.24, 1.02)	(-79, 211)	(-78, 201)	**0.649**	**0.664**
4	EXTERNAL COMB BOILERS-UTILITIES-COAL	1.124	28 (1)	L (-0.13, 1.51)	(-59, 123)	(-60, 130)	**0.576**	**0.610**
5	All Off-highway Vehicle: Diesel	0.339						
6	HAZARDOUS WASTE INCINERATION	0.276	48	L (1.80, 0.84)	(-26, 32)	(-28, 33)	**0.115**	**0.112**
7	HARD CHROMIUM ELECTROPLATING	0.230	12	W (0.58, 0.48)[h]	(-71, 137)	(-71, 137)	**0.356**	**0.206**
8	ORG. SOLV. EVAPORATION-SURF. COATG-GENL	0.181	10	L (1.39, 1.13)	(-60, 130)	(-61, 129)	0.059	0.076
9	CHROMIUM METAL PLATING	0.160			(-71, 137)	(-72, 130)	**0.356**	0.057
10	FABRICATED PLATE WORK (BOILER SHOPS)	0.120						
11	NONCLAY REFRACTORIES (NOT SUBJECT TO REFRACTORIES MANUFACTURING MACT)	0.120						
12a	Fuel Oil		13 (1)	L (1.91, 0.69)	(-32, 36)			
12b	Refinery gas and landfill gas		3	W (1.86, 2.48)	(-59, 69)			
12	OIL AND GAS FIELD MACHINERY MANUFACTURING	0.120			(-33, 34)	(-33, 34)	**0.099**	**0.104**

(Continued)

Table 3. Continued

Case No.	Emission Source Description	EI[a] (tons/yr)	n[b]	Variability in Emission Factor[c]	Uncertainty in Emission Factor (%, %)[d]	Uncertainty in Emission Inventory (%, %)[e]	Rank Correlation 1[f]	Rank Correlation 2[g]
13	Light Duty Gasoline Vehicles (LDGV)	0.107						
14	SECONDARY METAL PROD-STEEL FOUNDRIES	0.057	12	L (-0.21, 1.10)	(-55, 110)	(-54, 111)	-0.074	-0.056
15	ASPHALT ROOFING: DIPPING ONLY	0.053	5	L (1.83, 1.72)	(-91, 424)	(-91, 388)	**0.124**	**0.196**
16	Light Duty Gasoline Trucks 1 & 2 (LDGT)	0.046						
17	PORTLAND CEMENT MANUFACTURING	0.042						
18	MFG-VINYL ACETATE	0.030						
19	RESIDENTIAL HEATING: WOOD/WOOD RESIDUE	0.030	8	W (1.46, 0.86)	(-62, 108)	(-62, 107)	-0.043	-0.047
20	All Off-highway Vehicle: Gasoline, 2-Stroke	0.025						
21	All Off-highway Vehicle: Gasoline, 4-Stroke	0.024						
22	PRIMARY METAL PROD-STEEL PRODUCTION	0.023	3	L (0.74, 0.55)	(-49, 75)	(-49, 80)	-0.085	-0.079
23	RESIDENTIAL HEATING: DISTILLATE OIL	0.016						
24	EXTERNAL COMB BOILERS-COMML/INSTIT.-LIQ.	0.015						
25	INSTITUTIONAL/COMMERCIAL HEATING: DISTILLATE OIL	0.011						

(Continued)

Table 3. Continued

Case No.	Emission Source Description	EI[a] (tons/yr)	n[b]	Variability in Emission Factor[c]	Uncertainty in Emission Factor (%, %)[d]	Uncertainty in Emission Inventory (%, %)[e]	Rank Correlation 1[f]	Rank Correlation 2[g]
26	INDUSTRIAL BOILERS: RESIDUAL OIL	0.011	13 (1)	L (1.91, 0.70)	(-59, 69)	(-33, 38)	0.051	-0.016
27	PULP/PAPER IND.-KRAFT PULPING	0.010						

a. Point estimate of formaldehyde emission inventory
b. Sample size, for censored data set, the number of non-detects is shown in parenthesis
c. Variability in emission factor. L: lognormal distribution; G: gamma distribution; W: Weibull distribution. The parameters of the distribution are given in parenthesis
d. The 95 % confidence interval relative to the mean is given.
e. The 95% confidence interval relative to the mean is given. For the cases in which there is no information of variability in emission factors while there is information of uncertainty in the emission factor, the latter is based upon surrogate or previous work.
f. Rank correlation between the uncertainty in the total emissions and the uncertainty in the emission factors for each source category with correlated surrogates. Statistically significant correlations are shown in boldface
g. Rank correlation between the uncertainty in the total emissions and the uncertainty in the emission factors for each source category with uncorrelated surrogates. Statistically significant correlations are shown in boldface.
h. Rejected by the K-S test, but the fit is judged to be adequate

Table 4. Quantification of Variability and Uncertainty for Arsenic Emission Inventory

Case No.	Emission Source Description	EI[a] (tons/yr)	n[b]	Variability in Emission Factor[c]	Uncertainty in Emission Factor (%, %)[d]	Uncertainty in Emission Inventory (%, %)[e]	Rank Correlation 1[f]	Rank Correlation 2[g]
1	EXTERNAL COMB BOILERS-UTILITIES-COAL	1.77	29 (3)	L (-1.28, 2.34)	(-91, 264)	(-91, 272)	**0.986**	**0.985**
2	HAZARDOUS WASTE INCINERATION	0.347	45	L (0.63, 1.09)[h]	(-33, 52)	(-33, 50)	**0.105**	**0.103**
3	PORTLAND CEMENT MANUFACTURING	0.243						
4	PETROLEUM REFINERIES - CATALYTIC CRACKI	0.209						
5	Marine Vessels, Commercial	0.199						
6	PULP/PAPER IND.-KRAFT PULPING	0.0394						
7	RESIDENTIAL HEATING: DISTILLATE OIL	0.0275	3	W (2.63, 2.28)	(-51, 92)	(-52, 96)	-0.043	-0.049
8	RESIDENTIAL HEATING: WOOD/WOOD RESIDUE	0.0152			(-71, 71)	(-73, 76)	-0.046	-0.008
9	INDUSTRIAL BOILERS: RESIDUAL OIL	0.0136	13 (3)	W (1.03, 1.07)	(-46, 59)	(-46, 61)	**0.101**	**0.095**
10	INSTITUTIONAL/COMMERCIAL HEATING: DISTILLATE OIL	0.0123			(-51, 92)	(-53, 77)	-0.043	**-0.086**
11	INSTITUTIONAL/COMMERCIAL HEATING: RESIDUAL OIL	0.011			(-46, 59)	(-47, 71)	0.101	-0.003
12	INSTITUTIONAL/COMMERCIAL HEATING: BITUMINOUS AND LIGNITE	0.008			(-91, 264)	(-91, 261)	**0.986**	0.022
13	WOOD PRESERVING	0.005						

(Continued)

Table 4. Continued.

Case No.	Emission Source Description	EI[a] (tons/yr)	n[b]	Variability in Emission Factor[c]	Uncertainty in Emission Factor (%, %)[d]	Uncertainty in Emission Inventory (%, %)[e]	Rank Correlation 1[f]	Rank Correlation 2[g]
14	EXTERNAL COMB BOILERS-INDUSTRIAL-WOOD	0.00495	7 (2)	W (0.59, 0.61)	(-71, 71)	(-71, 73)	-0.045	-0.053
15	RESIDENTIAL HEATING: BITUMINOUS AND LIGNITE COAL	0.0044			(-91, 264)	(-91, 271)	**0.986**	**-0.121**
16	MFG-INORGANIC CHEMICALS-GENERAL PROCESSE	0.0034						
17	PRIMARY NONFERROUS METALS PRODUCTION	0.003						
18	INDUSTRIAL BOILERS: WOOD/WOOD RESIDUE	0.00266	7 (2)	W (0.59, 0.61)	(-71, 71)	(-72, 76)	0.058	0.058
19	INDUSTRIAL BOILERS: WASTE OIL	0.00222			(-46, 59)	(-45, 63)	**0.101**	0.022
20	FOOD AND AGRICULTURAL PRODUCTS: COTTON GINNING	0.00208						

a. Point estimate of formaldehyde emission inventory
b. Sample size, for censored data set, the number of non-detects is shown in parenthesis
c. Variability in emission factor. L: lognormal distribution; G: gamma distribution; W: Weibull distribution. The parameters of the distribution are given in parenthesis
d. The 95 % confidence interval relative to the mean is given.
e. The 95% confidence interval relative to the mean is given. For the cases in which there is no information of variability in emission factors while there is information of uncertainty in the emission factor, the latter is based upon surrogate or previous work.
f. Rank correlation between the uncertainty in the total emissions and the uncertainty in the emission factors for each source category with correlated surrogates. Statistically significant correlations are shown in boldface
g. Rank correlation between the uncertainty in the total emissions and the uncertainty in the emission factors for each source category with uncorrelated surrogates. Statistically significant correlations are shown in boldface.
h. Rejected by the K-S test, but the fit is judged to be adequate

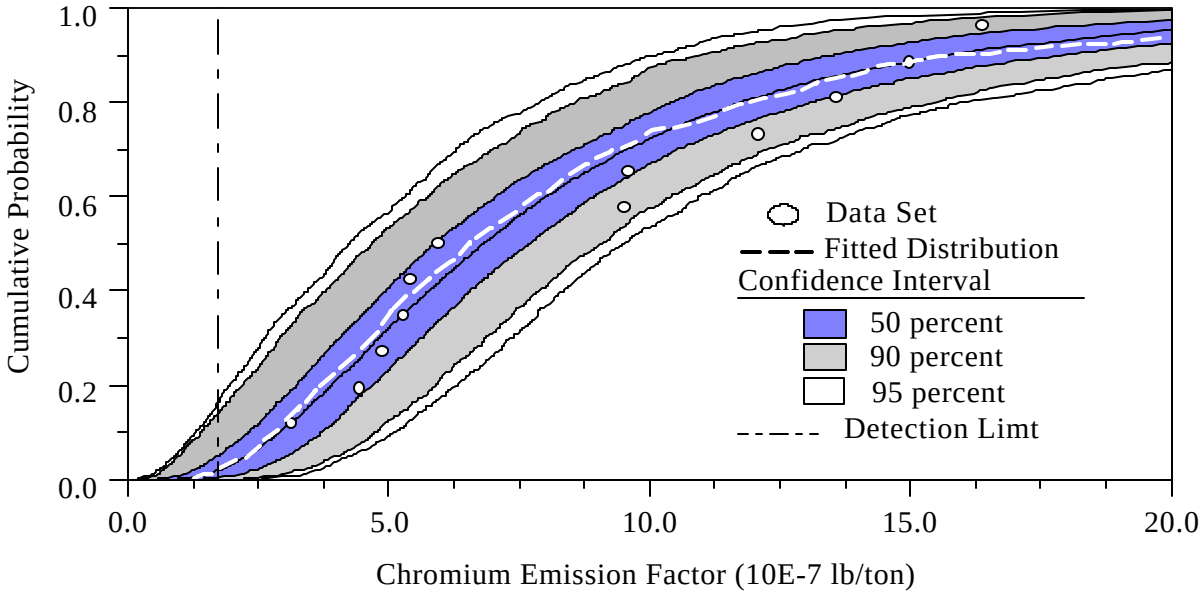


Figure 1. Variability and Uncertainty in Chromium Emission Factor for Case 26 (Industrial boilers: residual oil) Estimated Based Upon a Lognormal Distribution

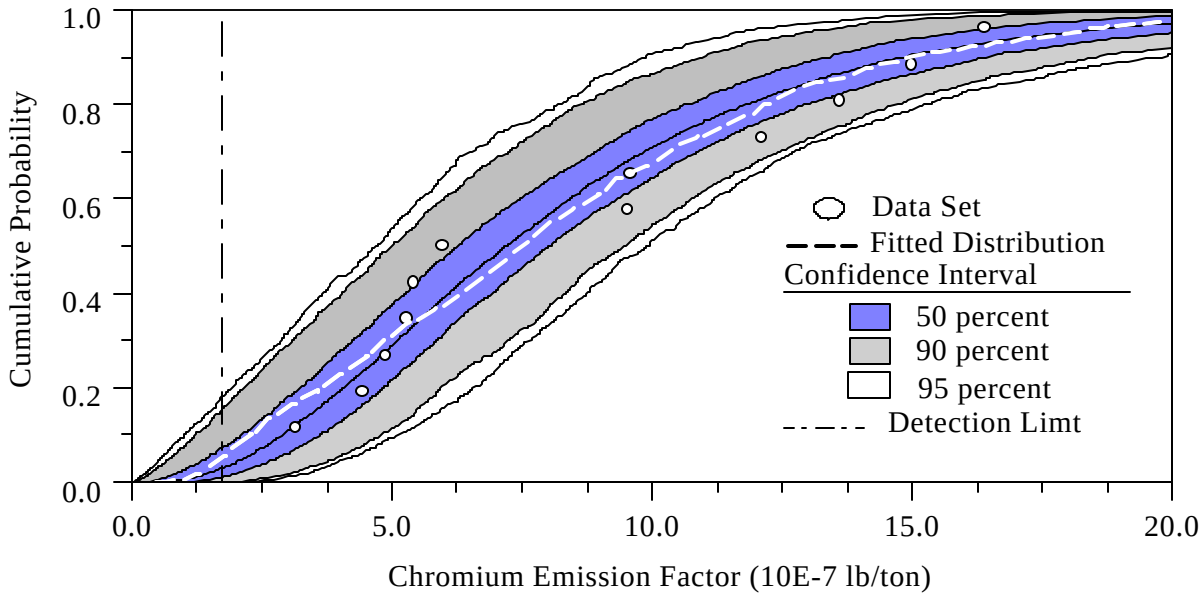


Figure 2. Variability and Uncertainty in Chromium Emission Factor for Case 26 (Industrial boilers: residual oil) Estimated Based Upon a Gamma Distribution

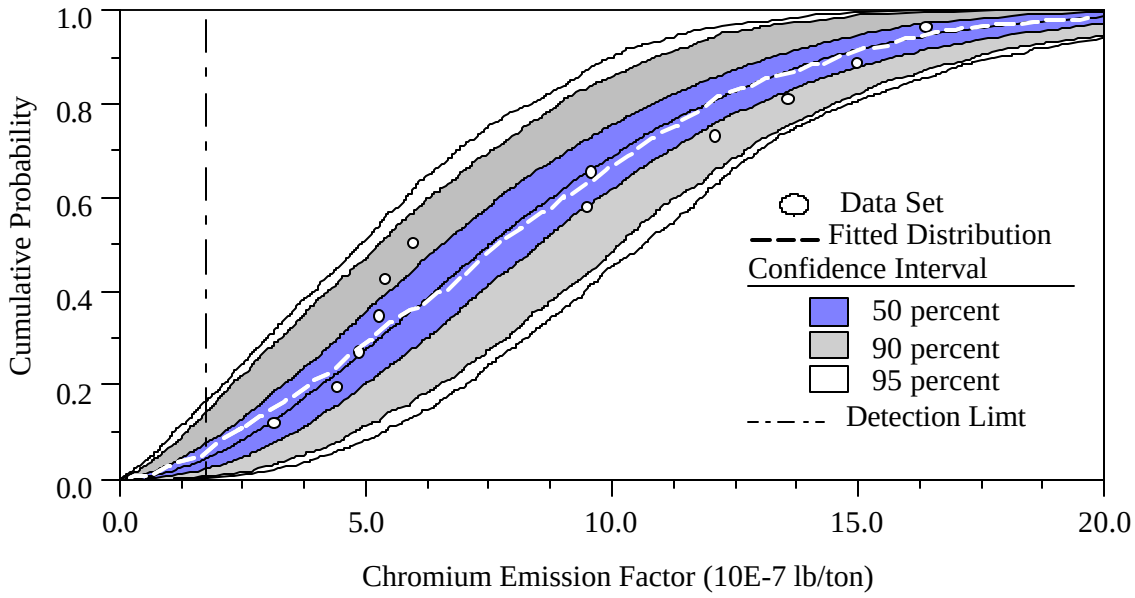


Figure 3. Variability and Uncertainty in Chromium Emission Factor for Case 26 (Industrial boilers: residual oil) Estimated Based Upon a Weibull Distribution

Supporting Information

The objective of this project is to develop probabilistic emission inventories (EI) of benzene, formaldehyde and chromium for the Houston area. A key step is to quantify variability. In this supporting information, the likelihood functions of Maximum Likelihood Estimation for censored data are given for lognormal, gamma and Weibull distributions. Empirical emission factor data and data source for benzene, formaldehyde, chromium and arsenic are shown in Tables S-1 to S-4. The method and results of quantification of uncertainty in emission factors for mobile sources and aggregation source categories are addressed. The uncertainties in the total emission inventories are given graphically for benzene, formaldehyde, chromium and arsenic.

Maximum Likelihood Estimation

The likelihood function for data without censoring is:

$$L(\theta_1, \theta_2, \ldots, \theta_k) = \prod_{i=1}^{n} f(x_i | \theta_1, \theta_2, \ldots \theta_k) \quad (1)$$

Where,

$\theta, \theta_2, \ldots, \theta_k$ = Parameters of the distribution

x_i = Values of random variable, for, i = 1, 2, …, n

n = Number of data points in the data set

f( ) = Probability density function

The likelihood function for censored data sets having multiple detection limits is: *(19, 38)*

$$L(\theta_1, \theta_2, \ldots \theta_3) = \prod_{i=1}^{n} f(x_i | \theta_1 \theta_2, \ldots, \theta_k) \left\{ \prod_{m=1}^{p} \left(\prod_{j=1}^{ND_m} F(DL_m | \theta_1, \theta_2, \ldots, \theta_k) \right) \right\} \quad (2)$$

Where,

$\theta, \theta_2, \ldots, \theta_k$ = Parameters of the distribution

x_i = Detected data point, where, i = 1, 2, …, n

ND_m = Number of non-detects corresponding to detection limit DL_m, where, m = 1, 2, …, P.

P = Number of detection limits

f( ) = Probability density function

F( ) = Cumulative distribution function

According to equation (2), for the lognormal distribution, the log-likelihood function including left-censored data is given by: *(22, 23)*

$$J(\mu,\sigma) = -n\ln\sigma - \frac{n}{2}\ln(2\pi) - \sum_{i=1}^{n}\left\{\frac{(x_i-\mu)^2}{2\sigma^2}\right\} + \sum_{m=1}^{P} ND_m \ln\left\{0.5\left[1+erf\left(\frac{DL_m-\mu}{\sigma\sqrt{2}}\right)\right]\right\} \quad (3)$$

For the gamma distribution, the log-likelihood function including left-censored data is given by: *(22, 23)*

$$J(\alpha,\beta) = -n\left\{\alpha\ln(\beta) + \ln[\Gamma(\alpha)]\right\} + \sum_{i=1}^{n}\left\{(\alpha-1)\ln(x_i) - \frac{x_i}{\beta}\right\} + \sum_{m=1}^{p} ND_m \ln\left\{\frac{\int_0^{DL_m/\beta} e^{-t}t^{\alpha-1}dt}{\Gamma(\alpha)}\right\} \quad (4)$$

For the Weibull distribution, the log-likelihoood function including left-censored data is given by: *(22, 23)*

$$J(\alpha,\beta) = -n\left(\frac{\alpha}{\beta}\right) + \sum_{i=1}^{n}\left\{(\alpha-1)\ln\left(\frac{x_i}{\beta}\right) - \left(\frac{x_i}{\beta}\right)^{\alpha}\right\} + \sum_{m=1}^{p} ND\ln\left\{1-\exp\left[-\left(\frac{DL}{\beta}\right)^{\alpha}\right]\right\} \quad (5)$$

Empirical Emission Factor Data

Empirical emission factor data of benzene, formaldehyde, chromium and arsenic are obtained. For some source categories, reasonable surrogate data are used. The data status to

indicate direct or surrogate data are in given in Tables S-1 to S-4 for each source category for the four pollutants. The references of the data sources are also given in Tables S-1 to S-4.

Quantification of Uncertainty in Emission Factors for Mobile Sources

For benzene and formaldehyde, data are available to quantify uncertainty in emission factors from mobile sources. For mobile source emission factors, previous work regarding uncertainty estimates has been done for total hydrocarbon emissions for both onroad and nonroad sources *(5, 7, 8)*. The uncertainty in the fraction of benzene and formaldehyde in total organic gas emissions has also been estimated for onroad LDGV sources *(37)*. Therefore, quantification of uncertainty in benzene and formaldehyde emission factors in gasoline mobile sources is based upon the results from previous work. For diesel mobile sources, the uncertainties in the THC emission factors and the fraction of benzene and formaldehyde in THC emissions were quantified based upon empirical data using bootstrap simulation. The uncertainty in benzene and formaldehyde emission factors was quantified based on the products of the uncertainties in THC emissions and percent of THC emitted as an air toxic. Here, the model used to quantify the uncertainty for benzene and formaldehyde from mobile sources is introduced.

The air toxic emission factors for gasoline mobile sources for either benzene or formaldehyde is given by:

$$EF_{toxic} = EF_{THC} \times F_{toxic/TOG} / 100 \times F_{TOG/THC} \quad (6)$$

Where,

EF_{toxic} = Toxic Emission factor (unit: g/mi for onroad mobile sources; g/hp-hr for nonroad lawn and garden engines; g/kWh for nonroad construction, farm and industrial engines)

EF_{THC} = THC emission factor (unit: g/mi for onroad mobile sources; g/hp-hr for nonroad lawn and garden engines; g/kWh for nonroad construction, farm and industrial engines)

$F_{toxic/TOG}$ = Toxic emission fraction (unit: % of TOG mass emitted as the selected air toxic)

$F_{TOG/THC}$ = Mass Ratio of TOG to THC emissions

The air toxic emission factors for diesel mobile sources for either benzene or formaldehyde is given by:

$$EF_{toxic} = EF_{THC} \times F_{toxic/THC} \quad (7)$$

Where,

$F_{toxic/THC}$ = Toxic mass emission fraction in THC emitted as the selected air toxic)

Most hydrocarbon emission data from mobile sources is measured as total hydrocarbon. THC is measured using a Flame Ionization Detector (FID) calibrated with propane *(SI-1)*. The FID is assumed to respond to all hydrocarbons identically as it responds to propane in determining the concentration of carbon atoms in a gas sample. Most hydrocarbons respond nearly identically as propane with notable exceptions being oxygenated hydrocarbons such as alcohols and aldehydes commonly found in engine exhaust. Because alcohols and especially aldehydes are chemically reactive and therefore ozone-forming hydrocarbons, the California Air Resources Board defined a measurement that adds the THC and oxygenated compounds into a new measurement called total organic gas. The oxygenated components are measured by collecting aldehydes on dinitro-phenylhydrazine impregnated filter traps and alcohols in chilled

water impingers. The aldehydes and alcohols are extracted and measured using chromatography to determine emission rates. The mass ratio of TOG/THC is approximately equal to 1 for mobile sources *(SI-1)*. EPA often uses the toxic fraction, developed as a percentage of the toxic compound of interested contained in TOG, to calculate the toxic emission estimates for mobile source. Previous work was done by Bammi to analyze the ratio of toxics emission in total TOG for LDGV *(37)*. Therefore, $F_{toxic/TOG}$ is used instead of $F_{toxic/THC}$ for gasoline mobile sources. For diesel mobile sources, empirical emission factor data for THC, benzene and formaldehyde were available. A similar procedure as for gasoline mobile sources was applied in order to quantify the uncertainty in the benzene and formaldehyde emission factors. However, since no data were reported by CRC for TOG from diesel mobile sources, $F_{toxic/THC}$ is used *(35)*.

Since the objective of this work is to quantify the relative uncertainty in emission factors, each input in Equations (6) and (7) was normalized to its respective mean value to obtain:

$$UF_{toxic} = UF_{THC} \times UF_{toxic/TOG} \times UF_{TOG/THC} \quad (8)$$

$$UF_{toxic} = UF_{THC} \times UF_{toxic/THC} \quad (9)$$

Where,

UF_{toxic} = Normalized uncertainty factor for EF_{toxic}

UF_{THC} = Normalized uncertainty factor for EF_{THC}

$UF_{toxic/TOG}$ = Normalized uncertainty factor for $F_{toxic/TOG}$

$UF_{TOG/THC}$ = Normalized uncertainty factor for $F_{TOG/THC}$

$UF_{toxic/THC}$ = Normalized uncertainty factor for $F_{toxic/THC}$

Figure S-1 illustrates the use of Equation (9) to calculate the uncertainty in benzene emission factors from onroad diesel mobile sources.

Quantification of Uncertainty in Aggregation Source Categories

Case 3 for benzene, Cases 1, 2, 3, 6 and 11 for formaldehyde and Case 12 for chromium are aggregations of several subcategories. For such categories, the uncertainty for the main source category was obtained based on the weighted average uncertainty from the subcategories. The weight assigned to each subcategory is based on the relationship between the subcategories. However, the information to determine the relationship is not available for most cases and thus assumptions were made.

For Case No. 3 of the petroleum refinery source category in Table 1, an approximate weight for each subcategory was defined based on the gasoline marketing distribution system in the United States *(33)*. The scheme of the distribution system is shown in Figure S-2. According to the scheme and the subcategories for which empirical data are available, including Case Nos. 3a to 3i in Table 1, Source No. 3 was taken as the combination of the following six subcategories:

1. Petroleum refinery storage tanks (PRST) represented by Case No. 3i;
2. Bulk terminal (BT) represented by Case Nos. 3a, 3b and 3c;
3. Typical bulk plant (TBP) represented by Case Nos. 3a and 3d;
4. Storage losses at typical pipeline breakout station (TPBS) represented by Case Nos. 3e and 3f;
5. Service station (TSS) represented by Case No. 3g; and
6. Emissions from wastewater (WW) represented by Case No. 3h.

For the subcategories of BT, TBP and TPBS, there are several subcomponents. First, the weights for the six subcategories were defined. Then, the weights were defined for the subcomponents in BT, TBP and TPBS.

According to the scheme of the distribution system, PRST and BT are on the main flow of the distribution system, and TBP, TPBS, and TSS are with respect to partial flow of the distribution system. The subcategories on the main flow of the distribution system were assigned weights of one unit, and the emission sources on the partial flow were assigned weights of one-half. Wastewater collection and treatment is not shown in Figure S-2 and there is no available information to determine the emissions of air toxics from WW relative to the total emissions from petroleum refinery. According to EPA, air emissions from petroleum refinery wastewater collection and treatment are one of the largest sources of VOC emissions at a refinery. *(33)* Therefore, for purpose of calculation, the WW subcategory was assigned a weight of one. The weighted uncertainty factor for Source 3 is calculated as:

$$UF_3 = (UF_{PRST} + UF_{BT} + 0.5 \times UF_{TBP} + 0.5 \times UF_{TBPS} + 0.5 \times UF_{TSS} + UF_{ww})/4.5 \quad (10)$$

Where, UF_3, UF_{PRST}, UF_{BT}, UF_{TBP}, UF_{TBPS}, UF_{TSS} and UF_{ww} are normalized uncertainty factors of benzene emission factors from Emission Sources 3, PRST, BT, TBP, TBPS, TSS and WW, respectively.

Since subcategories BT, TBP and TBPS are each composed of several subcomponents, a weight was needed for each subcomponent. For components that represent different processes, equal weights were assigned to each component. For example, weights of 0.5 were assigned to subcomponents 3a and 3d in subcategory TBP. If the components represent different seasons, a weight of ¾ was assigned to the non-winter season and ¼ was assigned to winter season on the assumption that winter corresponds to three winter months per year. For example, a weight of ¾ was assigned to 3e and ¼ was assigned to 3f in TBPS. For BT, there are two levels of subcomponents. In the first level, gasoline loading racks at bulk terminal, 3a, was one subcomponent and was assigned a weight of ½. Storage losses at gasoline bulk terminal

(SLGBT) composed of Case Nos. 3b and 3c was other subcomponent and was assigned a weight of 1/2. In the second level, weights were assigned to 3b and 3c in SLGBT. Since 3b is for the nonwinter season, it was assigned a weight of 3/4. Similarly, since 3c is for the winter season, it was assigned a weight of 1/4.

Based on the weighted average, the 95 percent confidence interval of the mean emission factors for Source No. 3 was obtained as minus 55 percent to plus 158 percent. This is less than the range of uncertainty for most of the subcategories. The reason is that there is no correlation of the uncertainty factors among the subcategories and the uncertainties from the subcategories tend to offset each other when using Equation (10). For comparison, uncertainty was also estimated for a straight arithmetic average of all subcategories. The resulting 95 percent relative confidence interval of the mean is from minus 53 percent to plus 127 percent. Although there is some difference in the two results at the upper level of the range, the two results are comparable in magnitude. Therefore, although the method for determining the weight of each subcategory is approximate, the final results are not strongly sensitive to the weights for this case.

For formaldehyde mobile sources, such as Source Nos. 1 and 2 in Table 2, the emission of each subcategory from Houston is available except regarding the breakdown values of 2-stroke and 4-stroke nonroad gasoline lawn and garden mobile sources as shown in Table S-2. In these two cases, the emission fraction of each subcategory is assigned as the weight to calculate the weighted average uncertainty of the aggregation category. For 2-stroke and 4-stroke nonroad gasoline lawn and garden mobile sources, the formaldehyde emission fractions are available from Jacksonville, Florida and are used here for the weights. For example, for Source No. 1 in Table 6, the ratio of the emissions from 1a to 1b is 5:27 according to Jacksonville emission inventory *(SI-2)*. Therefore, the weight for 1a is assigned to be 5/(5+27), which is 0.156, and the

weight for 1b is assigned to be 0.844 to get the weighted average uncertainty for gasoline nonroad mobile sources. The weights of gasoline nonroad mobile source, diesel nonroad mobile source and aircraft source are assigned according to the emissions from Houston. That is, the weight of gasoline nonroad mobile source is equal to 182.79/1281.54, which is 0.143. The weights of diesel nonroad mobile sources and aircraft are equal to 0.729 and 0.128, respectively.

For Source Nos. 3, 6 and 11 in Table 2, there is no breakdown available regarding the formaldehyde emissions for each subcategory in Houston. As an approximate assumption, equal weights were assigned to each subcategory within these three major categories. That is, for Source Nos. 3 and 6, it is assumed that liquid fuel and gas fuel contribute equally to uncertainty. For Source No. 11, it is assumed that coal combustion, wood fired waste combustion and fuel oil combustion contribute equally to the uncertainty in the emissions. For Source No. 3, the uncertainty in subcategories 3a and 3b are relatively large compared to other source categories; thus, the resulting weighed average uncertainty for Source No. 3 will still be relatively large no matter what weights are assigned for each subcategory. Therefore, the uncertainty results are insensitive to the choice of a weight and are robust in the absence of data regarding what the weight should be. For Source Nos. 6 and 11, from the sensitivity study, the uncertainty in the total emissions has a small correlation to the emission factor uncertainty in these two cases. Therefore, the uncertainties associated with weights for these cases are not important to the uncertainty in total emissions.

For Source No. 12 of Chromium emission factors from oil and gas fired machinery manufacturing, there is no breakdown available for oil and gas fired sources from Houston, the fractions of the emissions for oil industry combustion source and gas industry combustion source from Jacksonville are used as the weights, which are 0.55 and 0.45, respectively (*SI-2*).

Based on the approaches described here, the 95 percent confidence intervals of the mean of benzene, formaldehyde and chromium emission factors were obtained for the source categories composed of several subcategories. The results are shown in Tables 1, 2 and 3.

Uncertainty in Total Emission Inventories

The uncertainties in the total urban air toxics emissions are shown by the cumulative probability distributions of the normalized uncertainty factors in Figures S-3 to S-10 for the four pollutants. The results based on both correlated surrogates and uncorrelated surrogates are given.

REFERENCES

1. Lindhjem, C.E. (1997), "Conversion Factors for Hydrocarbon Emission Components," Assessment and Modeling Division, U.S. Environmental Agency Office of Mobile Sources. http://www.epa.gov/otaq/regs/toxics/r99029.pdf

2. Tilley, L. (2003), Personal communication with Lori Tilley, Regulatory and Environmental Services Department, City of Jacksonville, FL, via email, May 26th, 2003.

Table S-1. Data and Data Source for Benzene Emission Factors

Case No.	Emission Source Description	Data Status[a]	Data Source
1	Mobile Source – LDGV	D	References: 7 & 35
2	Mobile Source – LDGT	S	Surrogate: Case 1
3a	Gasoline loading racks at bulk terminals and bulk plants	D	References: 30 & 31
3b	Storage losses at a typical gasoline bulk terminal (non Winter)	D	
3c	Storage losses at a typical gasoline bulk terminal (winter)	D	
3d	For a typical bulk plant	D	
3e	Storage losses at a typical pipeline breakout station (non-winter)	D	
3f	Storage losses at a typical pipeline breakout station (winter)	D	
3g	For typical Service Station for petroleum refinery	D	
3h	Petroleum Refinery wastewater	D	
3i	Storage tank for petroleum refinery	D	
3	Weighted average from 3a to 3i	D	
4	4-stroke lawn and garden engines	D/S	References: 3 & 35
5	2-stroke lawn and garden engines	D/S	References: 3 & 35
6	Construction, farm and industrial engine (diesel 4S)	D/S	References: 32 & 33
7	Oil and natural gas production	S	Surrogate: methane fugitive emissions from gas and oil industry Reference: 34
8	Storage and transport, Natural Gas Transmissions and Marine Transport	S	Surrogate: methane fugitive emissions from gas and oil industry Reference: 34
9	Mobile source-HDGV	S	Surrogate: Case 1
10	Other combustion-forest wildfires	D	Reference: 30
11	Solid waste disposal- sewage treatment	D	Reference: 30
12	Industrial Processes; Chemical Manufacturing; Acetylene production	S	Reference: 34
13	Fuel oil external combustion	D	Reference: 28
14	Typical ethylene plant	D	Reference: 30
15	Storage and Transport; Petroleum Product Storage; Gasoline Service Stations;	D	Reference: 30

(Continued)

Table S-1. Continued.

Case No.	Emission Source Description	Data Status[a]	Data Source
16	Industrial Processes; Petroleum Industry; Fugitive	S	Surrogate: methane fugitive emissions from gas and oil industry Reference: 34
17	Other combustion-managed prescribed burning	D	Reference: 30
18-1	THC from HDDV	D	Reference: 32
18-2	Benzene fraction in THC from HDDV	D	Reference: 32
18	Benzene emission factor from HDDV	D	Reference: 32
19	Industrial Processes; Chemical Manufacturing; Fugitive Emissions	S	Surrogate: methane fugitive emissions from gas and oil industry Reference: 34
20	Mobile source-aircraft	D/S	Reference: 35
21	Industrial Processes; Petrolem Industry; Fugitive Emissions; Miscellaneous: Sampling/Non-Asphalt	S	Surrogate: methane fugitive emissions from gas and oil industry Reference: 34
22	Petroleum refinery-process vent in refinery product	D	Reference: 32
23	Loading, ballasting and transit losses from marine vessels	D	Reference: 32
24	Industrial Processes; Chemical Manufacturing; Processes; Fugitive leaks	S	Reference: 34

[a] D = directly relevant data; S = surrogate data; D/S = Directly relevant data for THC, surrogate data for % of TOG (or THC) emitted as benzene.

Table S-2. Data and Data Source for Formaldehyde Emission Factors

Case No.	Emission Source Description	Data Status[a]	Data Source
1a	Nonroad - 4-stroke lawn and garden engines	D/S	Reference: 3 & 35
1b	Nonroad - 2-stroke lawn and garden engines	D/S	Reference: 3 & 35
1c	Nonroad-CFI engine (diesel 4S)	D/S	Reference: 32 & 33
1d	Nonroad-Aircraft	D/S	Reference: 35
1	Nonroad	D/S	Reference: 3, 32, 33 and 35
2a	Onroad gasoline	D	Reference: 7 & 35
2b-1	Onroad diesel engines-THC	D	Reference: 32
2b-2	Onroad diesel engines- formaldehyde fraction	D	Reference: 32
2b	Onroad diesel engines	D	Reference: 32
2	Onroad	D	Reference: 7, 32 and 35
3a	Stationary reciprocating internal combustion engines (material type: liquid)	D	Reference: 31
3b	Stationary reciprocating internal combustion engines (material type: gas)	D	Reference: 31
3	Internal combustion engines	D	Reference: 31
4	Oil and gas extraction	S	Surrogate: methane fugitive emissions from gas and oil industry Reference: 34
5	Chemical and allied processes	S	Surrogate: methane fugitive emissions from gas and oil industry Reference: 34
6a	Stationary combustion turbines (material type: liquid)	D	Reference: 31
6b	Stationary combustion turbines (material type: gas)	D	Reference: 31
6	Weighted average of 6a and 6b	D	Reference: 31
7	Petroleum refineries	D	Reference: 26
8	Open burning, forest and wildfires	---	---
9	Open burning, prescribed burnings	---	---
10	Utility boilers	S	Surrogate: Case 11
11a	External combustion- coal combustion	D	Reference: 27

(Continued)

Table S-2. Continued.

Case No.	Emission Source Description	Data Status[a]	Data Source
11b	External combustion-wood fired waste	D	Reference: 29
11c	External combustion-fuel oil	D	Reference: 28
11	Industry boilers	D	Reference: 27, 28 and 29
12	Structure fires	---	---

[a] D = directly relevant data; S = surrogate data; D/S = Directly relevant data for THC, surrogate data for % of TOG (or THC) emitted as formaldehyde.

Table S-3. Data and Data Source for Chromium Emission Factors

Case No.	Emission Source Description	Data Status[a]	Data Source
1	PETROLEUM REFINERIES - CATALYTIC CRACKI	D	Reference: 31
2	Marine Vessels, Commercial		
3	CHEM MFG-FUEL FIRED EQ-PROCESS HEATERS	D	Reference: 37
4	EXTERNAL COMB BOILERS-UTILITIES-COAL	S	Surrogate: industrial boilers Reference: 27
5	All Off-highway Vehicle: Diesel		
6	HAZARDOUS WASTE INCINERATION	D	Reference: 36
7	HARD CHROMIUM ELECTROPLATING	D	Reference: 31
8	ORG. SOLV. EVAPORATION-SURF. COATG-GENL	D	Reference: 31
9	CHROMIUM METAL PLATING	S	Surrogate: Case7
10	FABRICATED PLATE WORK (BOILER SHOPS)		
11	NONCLAY REFRACTORIES (NOT SUBJECT TO REFRACTORIES MANUFACTURING MACT)		
12a	Fuel Oil		Reference: 28
12b	Refinery gas and landfill gas		Reference: 31
12	OIL AND GAS FIELD MACHINERY MANUFACTURING	S	Reference: 28 and 31
13	Light Duty Gasoline Vehicles (LDGV)		
14	SECONDARY METAL PROD-STEEL FOUNDRIES	D	Reference: 31
15	ASPHALT ROOFING: DIPPING ONLY	S	Surrogate: asphalt concrete Reference: 31
16	Light Duty Gasoline Trucks 1 & 2 (LDGT)		
17	PORTLAND CEMENT MANUFACTURING		
18	MFG-VINYL ACETATE		
19	RESIDENTIAL HEATING: WOOD/WOOD RESIDUE	S	Surrogate: boilers Reference: 29
20	All Off-highway Vehicle: Gasoline, 2-Stroke		
21	All Off-highway Vehicle: Gasoline, 4-Stroke		

(Continued)

Table S-3. Continued.

Case No.	Emission Source Description	Data Status[a]	Data Source
22	PRIMARY METAL PROD-STEEL PRODUCTION	D	Reference: 31
23	RESIDENTIAL HEATING: DISTILLATE OIL		
24	EXTERNAL COMB BOILERS-COMML/INSTIT.-LIQ.		
25	INSTITUTIONAL/COMMERCIAL HEATING: DISTILLATE OIL		
26	INDUSTRIAL BOILERS: RESIDUAL OIL	D	Reference: 28
27	PULP/PAPER IND.-KRAFT PULPING		

[a] D = directly relevant data; S = surrogate data;

Table S-4. Data and Data Source for Arsenic Emission Factors

Case No.	Emission Source Description	Data Status[a]	Data Source
1	EXTERNAL COMB BOILERS-UTILITIES-COAL	D	Reference: 27
2	HAZARDOUS WASTE INCINERATION	D	Reference: 36
3	PORTLAND CEMENT MANUFACTURING	D	Reference: 31
4	PETROLEUM REFINERIES - CATALYTIC CRACKI		
5	Marine Vessels, Commercial		
6	PULP/PAPER IND.-KRAFT PULPING		
7	RESIDENTIAL HEATING: DISTILLATE OIL	S	Surrogate: distillate oil turbine Reference: 31
8	RESIDENTIAL HEATING: WOOD/WOOD RESIDUE	S	Surrogate: Case 14
9	INDUSTRIAL BOILERS: RESIDUAL OIL	D	Reference: 28
10	INSTITUTIONAL/COMMERCIAL HEATING: DISTILLATE OIL	S	Surrogate: Case 7 Reference: 28
11	INSTITUTIONAL/COMMERCIAL HEATING: RESIDUAL OIL	S	Surrogate: Case 9
12	INSTITUTIONAL/COMMERCIAL HEATING: BITUMINOUS AND LIGNITE	S	Surrogate: Case 1
13	WOOD PRESERVING		
14	EXTERNAL COMB BOILERS-INDUSTRIAL-WOOD	D	Reference: 29
15	RESIDENTIAL HEATING: BITUMINOUS AND LIGNITE COAL	S	Surrogate: Case 1
16	MFG-INORGANIC CHEMICALS-GENERAL PROCESSE		
17	PRIMARY NONFERROUS METALS PRODUCTION		
18	INDUSTRIAL BOILERS: WOOD/WOOD RESIDUE (area source)	D	Reference: 29
19	INDUSTRIAL BOILERS: WASTE OIL	S	Surrogate: Case 9
20	FOOD AND AGRICULTURAL PRODUCTS: COTTON GINNING		

[a] D = directly relevant data; S = surrogate data;

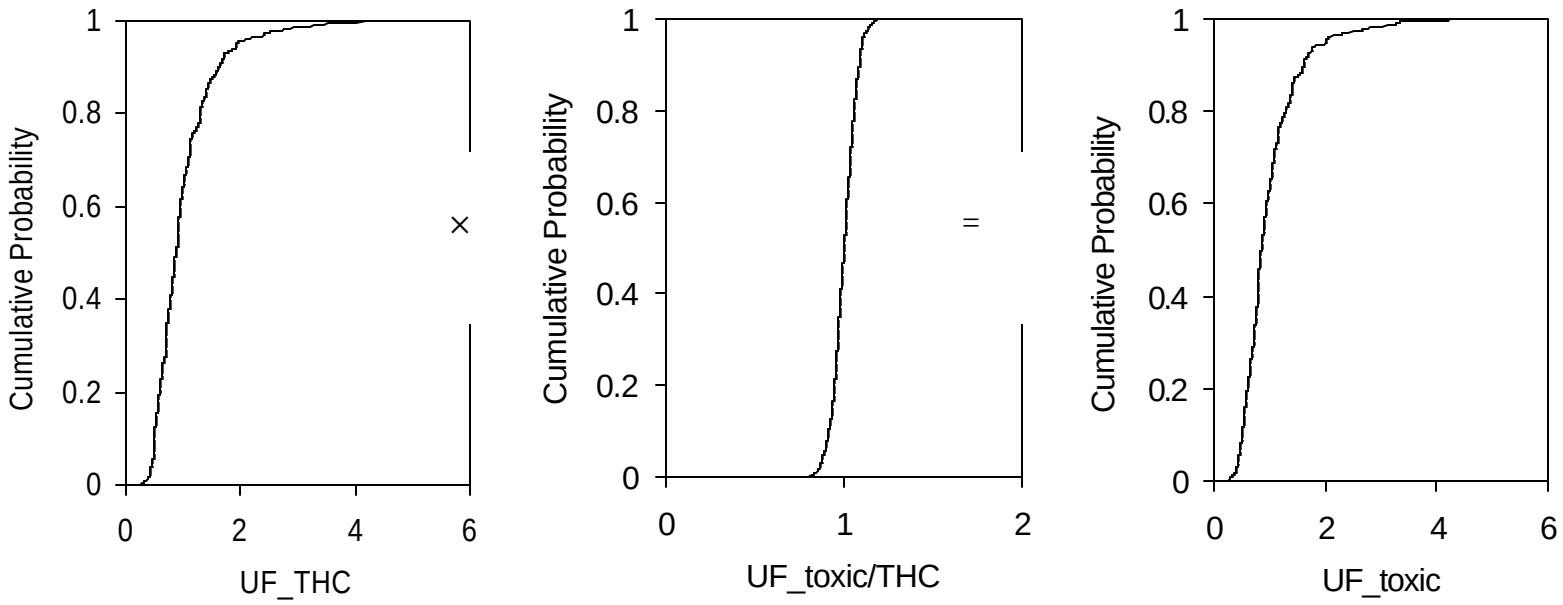


Figure S-1. Quantification of Uncertainty in the Mean of the Benzene Emissions Factor for Onroad Diesel Mobile Sources

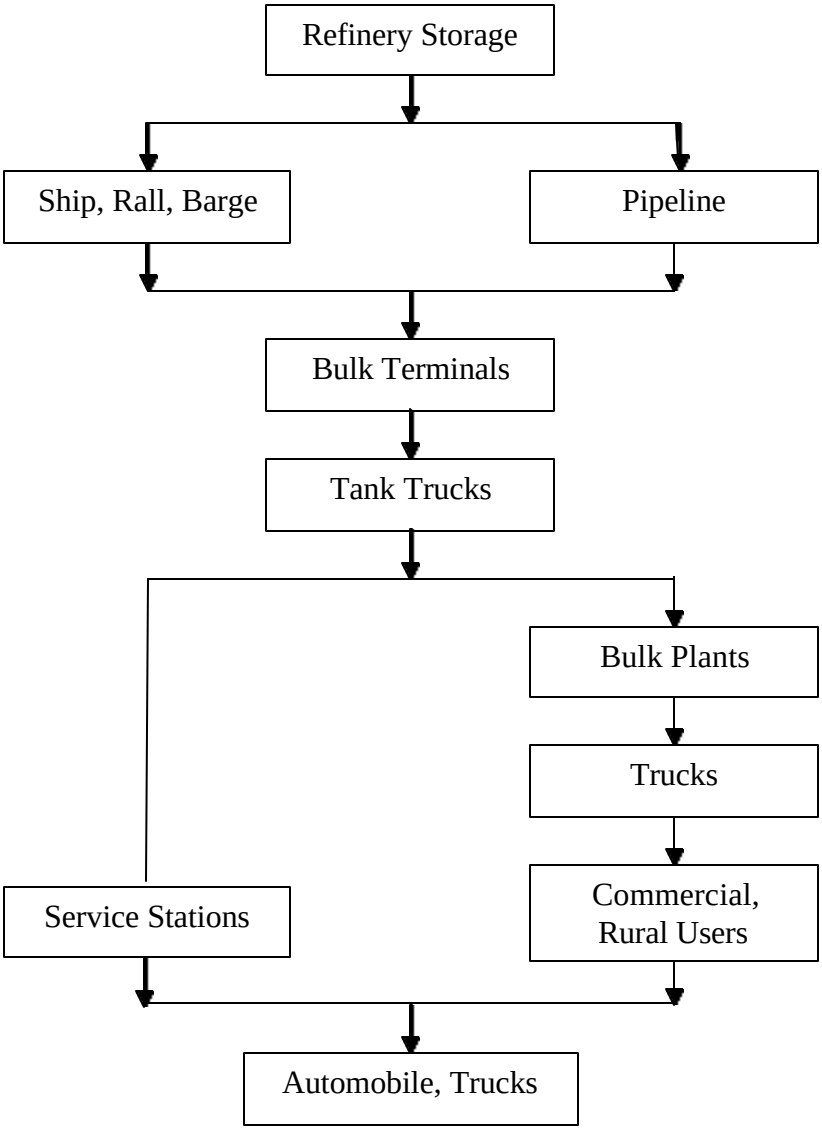


Figure S-2. The Gasoline Marketing Distribution System in the United States

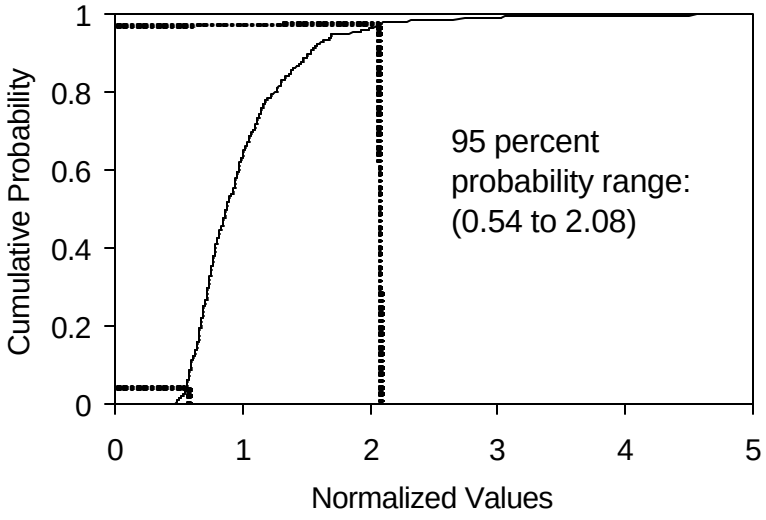


Figure S-3. Weighted Average Uncertainty Factor for Benzene Emissions for All Source Categories with Correlated Surrogates

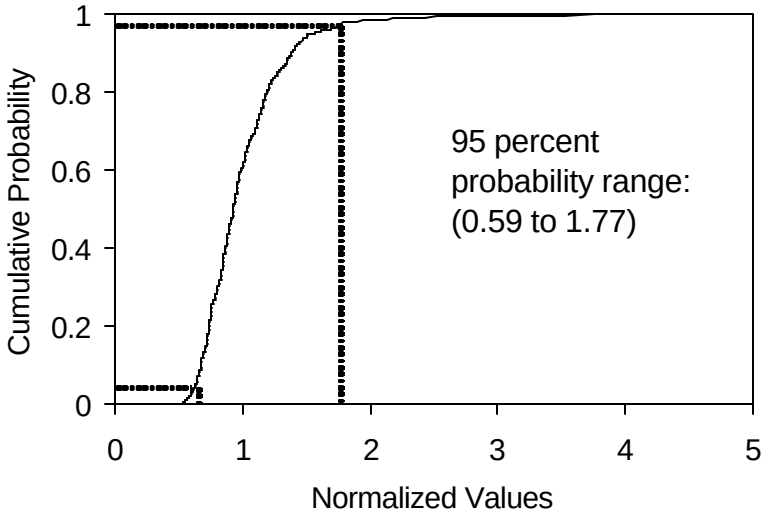


Figure S-4. Weighted Average Uncertainty Factor for Benzene Emissions for All Source Categories with Uncorrelated Surrogates

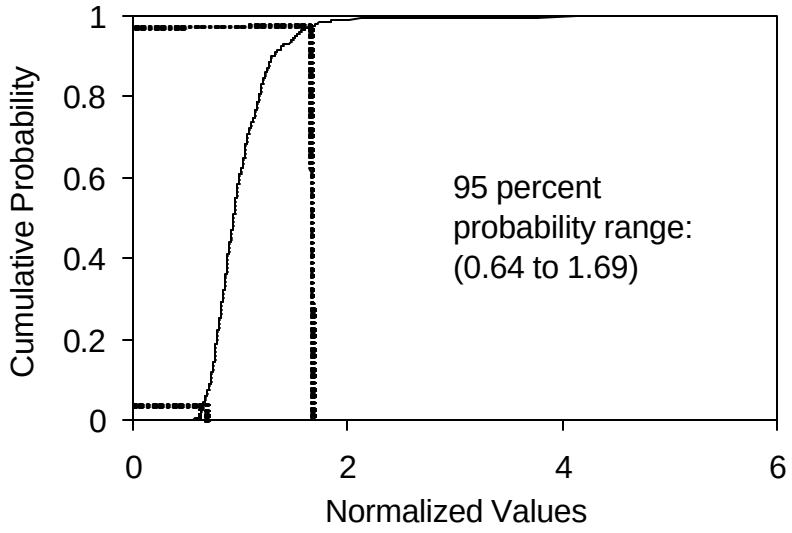


Figure S-5. Weighted Average Uncertainty Factor for Formaldehyde Emissions for All Source Categories with Correlated Surrogates

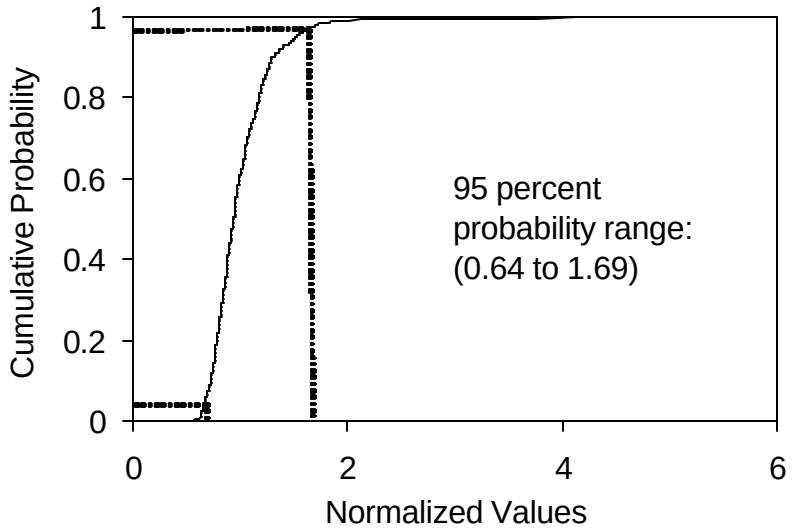


Figure S-6. Weighted Average Uncertainty Factor for Formaldehyde Emissions for All Source Categories with Uncorrelated Surrogates

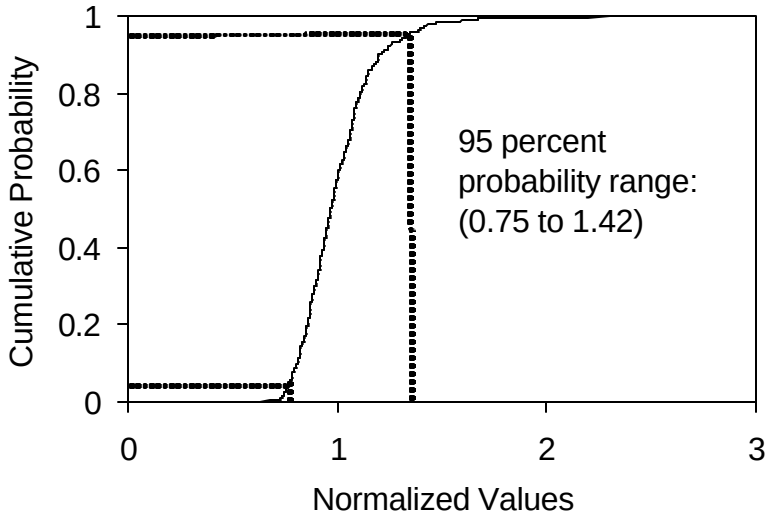


Figure S-7. Weighted Average Uncertainty Factor for Chromium Emissions for All Source Categories with Correlated Surrogates

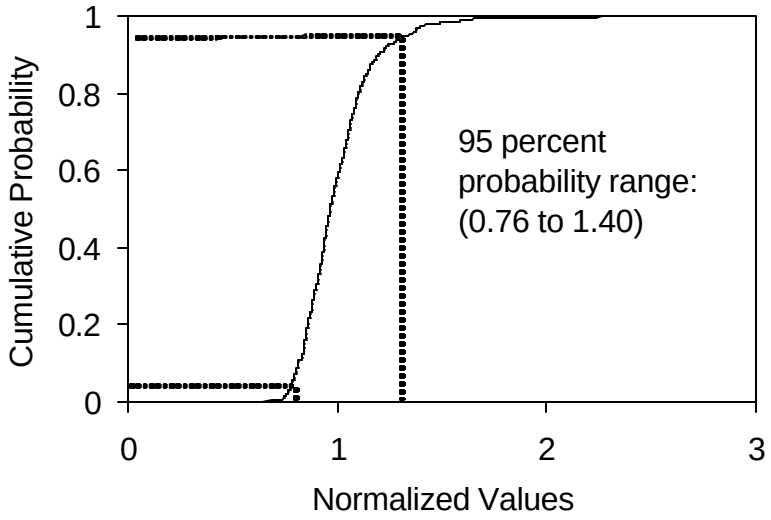


Figure S-8. Weighted Average Uncertainty Factor for Chromium Emissions for All Source Categories with Uncorrelated Surrogates

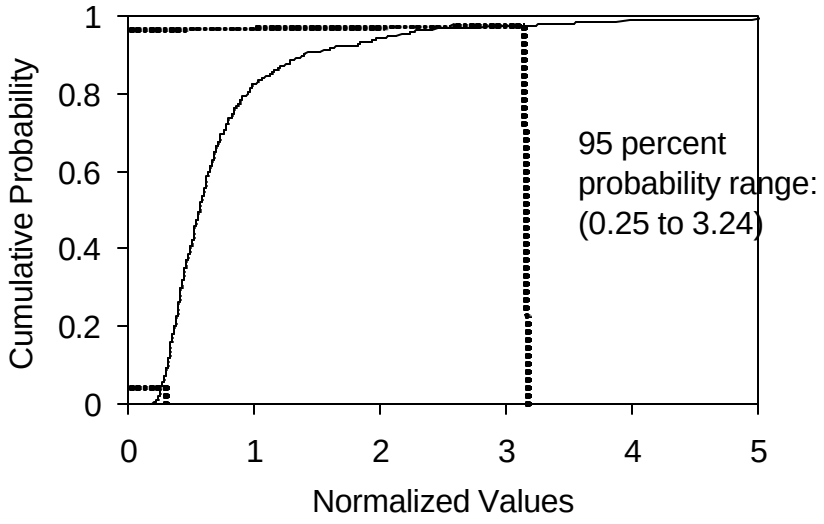


Figure S-9. Weighted Average Uncertainty Factor for Arsenic Emissions for All Source Categories with Correlated Surrogates

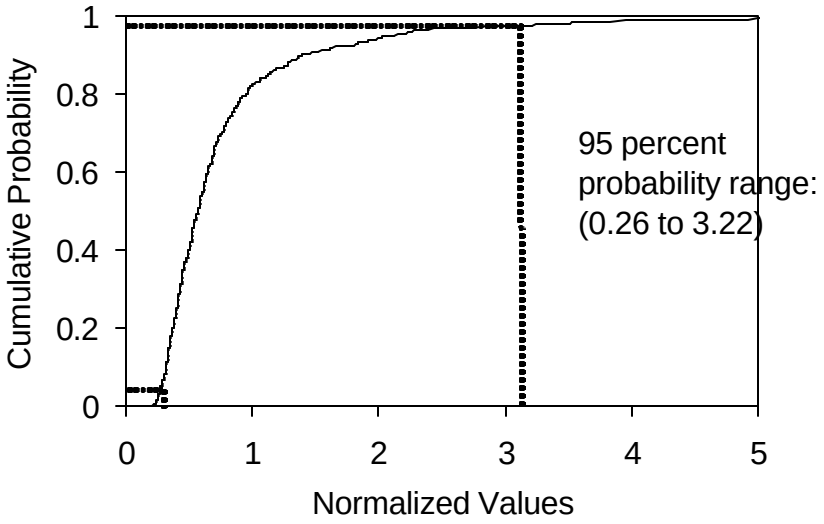


Figure S-10. Weighted Average Uncertainty Factor for Arsenic Emissions for All Source Categories with Uncorrelated Surrogates

Part V

DEVELOPMENT OF PROBABILISTIC EMISSION INVENTOTY OF AIR TOXICS FOR JACKSONVILLE, FLORIDA

Prepared to be submitted to

Journal of Air and Waste Management Association

Abstract. Probabilistic emission inventories were developed for 1, 3-butadiene, mercury, arsenic, benzene, formaldehyde and lead for Jacksonville, Florida. Maximum Likelihood Estimation (MLE) or Method of Matching Moments (MOMM) is used to fit parametric distributions to empirical emission factor data to represent inter-unit variability. Parametric bootstrap simulation and empirical bootstrap simulation are respectively applied to uncensored and censored data to quantify the uncertainty in urban air toxics emission factors. The probabilistic emission inventories were developed based on the product of the uncertainties in the emission factors and in the activity factors. The uncertainties in the urban air toxics emission inventories are typically large. For example, the 95 percent confidence interval ranges from minus 83 percent to plus 243 percent for arsenic. The key sources of uncertainty in the emission inventory for each toxic are identified by sensitivity study.

Key Words: Probabilistic emission inventory, Urban air toxics, 1,3-butadiene, Mercury, Arsenic, Benzene, Formaldehyde, Lead, Maximum likelihood estimation, Bootstrap simulation

1.0 Introduction

Quantification of uncertainty in emissions factors and emission inventories (EIs) is increasingly recognized as a need and there are a growing number of examples of such efforts. The National Research Council has repeatedly recommended that uncertainty in emissions be quantified.[1-3] The Intergovernmental Panel on Climate Change (IPCC) has developed good practice guidelines for quantification of uncertainty in greenhouse gas emissions estimates.[4] The U.S. Environmental Protection Agency has developed guidance regarding methods for uncertainty analysis. Uncertainties have been assessed quantitatively for emission factors, including source categories such as power plants, wood furniture coating, onroad mobile,

nonroad mobile, and natural gas engines, primarily with regard to ozone precursors or greenhouse gases.[5-16] Uncertainty has been quantified for selected emission inventories,[17-20] although in several cases simplifying assumptions were made regarding normality.[17, 19] Methods for distinguishing between variability and uncertainty and for dealing with various cases of practical significance, such as mixtures of distributions and data that contain non-detects, have been developed.[21-24]

A critical need for uncertainty analysis is with respect to urban air toxic emissions. EPA has developed a list of 33 urban air toxics, that represent the priority for additional assessment of the health effects of air toxics in urban areas.[25] There is a need to develop emission inventories of such pollutants for individual urban areas and to perform exposure and risk analysis with regard to human health effects. The National Research Council has strongly recommended a probabilistic approach to quantification of variability and uncertainty in exposure assessment,[2] and EPA has responded with guidelines on Monte Carlo analysis.[26] Therefore, there is a need to develop probabilistic emission inventories of urban air toxics in order to support probabilistic exposure assessment.

The purpose of this paper is to demonstrate methods for the development of probabilistic emission factors and inventories based upon case studies for selected urban air toxics for a specific urban area. The following pollutants were selected based upon consideration of their priority and data availability: 1, 3-butadiene, mercury, benzene and formaldehyde are ranked in the top 10 of EPA's list of 33 urban air toxics, and mercury, arsenic and lead are included in the list and are acute toxic agents of significant environmental and public health interest.[27, 28] A key challenge in dealing with air toxics emissions data is that many of the attempted measurement results are reported as below a detection limit. Therefore, there is a need to apply rigorous

statistical methods for dealing with non-detects in the process of quantification of inter-unit variability and uncertainty in the mean for emission factors of individual source categories. Among urban areas, Jacksonville, Florida has developed an extensive deterministic air toxic emission inventory, including 107 pollutants using 2000 as a base year.[29] The Jacksonville inventory is selected as the basis for case studies demonstrating the development of probabilistic emission inventories for the six selected urban air toxics.

The specific objectives of this paper are to: (1) quantify variability and uncertainty in air toxics emission factors for the largest emission sources of selected urban air toxics; (2) to develop probabilistic emission inventories for a specific urban area; (3) to identify key sources of uncertainty in the emission inventories for purposes of prioritizing future data collection.

2.0 Methodology

In this section, the candidate parametric distribution types used to represent the inter-unit variability of the emission factors are introduced. Then the methods to quantify the variability and uncertainty in the emission factors and to guide the choice of the preferred distribution, including the maximum likelihood estimation (MLE) and the method of matching moments (MOMM), bootstrap simulation, and goodness-of-fit test are addressed respectively. Finally, the method regarding developing probabilistic emission inventories and sensitivity study to identify key sources of uncertainty are addressed.

Because of inherent limitations of sampling and analytical chemistry measurement methods, urban air toxics data often contain several observations reported as below detection limits, referred to as "censored".[30] These data sets can have multiple detection limits. Multiple detection limits arise when individual measurements are collected by different sampling and

analytical procedures at different facilities within a source category. How to apply the methods to censored emission factor data was also discussed in this section.

2.1 Candidate Parametric Distributions for Inter-Unit Variability in Emission Factors

For environmental data sets, such as for concentrations or emission factors, lognormal, gamma and Weibull distributions are often chosen as parametric distributions to represent variability in data.[5, 11, 31-33] One of the most widely used distributional forms in probabilistic assessment is the lognormal distribution. The lognormal distribution describes random variables resulting from multiplicative processes.[34, 35] The gamma distribution is non-negative, positively skewed, and similar to the lognormal distribution in many cases but it is less "tail heavy". The Weibull distribution is a flexible distribution that can assume negatively skewed, symmetric, or positively skewed shapes.[31] It may be used to represent non-negative quantities. Since the emission factors are often non-negative and positively skewed, it is typically not reasonable to use normal distributions to describe variability in the emission factors. The lognormal, gamma and Weibull distributions are selected here as candidate distributions.

2.2 Maximum Likelihood Estimation and Method of Matching Moment

Two of the most common approaches to estimating the parameters of a distribution are MLE and MOMM.[31] MLE involves selecting values of the distribution parameters that are most likely to yield the observed data set. MLE can also be applied to fit distributions to censored data taking account of the non detect part.[24] MLE is asymptotically unbiased for both types of the data sets. MLE is used to fit distributions to all censored and most of the uncensored emission factor data sets. In MOMM, the parameters of a probability distribution model are selected so that the moments (e.g., mean, variance) of the model match the moments of the data set. MOMM is used to estimate distribution parameters for uncensored data when it gives a

better result than MLE in that it results in a smaller K-S test value. The estimated parametric distribution represents the inter-unit variability of the emission factor data.

2.3 Bootstrap Simulation

To quantify the uncertainty in the inter-unit variability and in selected statistics, such as the mean, for emission factors, conventional parametric bootstrap simulation is applied to uncensored data and empirical bootstrap simulation is applied to censored data.[24] In conventional parametric bootstrap simulation, a parametric probability distribution is fit to the original data set, which has a sample size of n. Monte Carlo simulation is used to randomly simulate synthetic data sets, referred to as bootstrap samples, each of sample size n. Typically, B bootstrap samples are simulated. For each bootstrap sample, a replication of a given statistic and its parametric distribution are estimated. Therefore, the confidence intervals for the given statistics or for the population distribution can be obtained based on the B replications.[11, 32, 36]

In the case of a censored data set, the conventional approach to parametric bootstrap simulation cannot be directly applied. Specifically, it is necessary to generate bootstrap samples so that there can be random variation in the number of data points that are below detection limit. In order to do this, an empirical bootstrap approach is used. In empirical bootstrap simulation, each of the original n data points is sampled with replacement and with equal probability of being sampled. In the original data set, either the value of data point is given for detected data or the detection limit is given for censored data. Therefore, for each data point, an indicator symbol d is given to indicate whether it is a detected value or below a detection limit. A value of d equal to 1 was used to represent a data point below a detection limit and d equal to 0 was used to represent a detected data point. In the case of non-detected data, the numerical value of the data point used in the bootstrap simulation was the detection limit itself. When generating bootstrap

samples from the original censored data set, both the data point value and its indicator symbol were sampled together. Therefore, for each bootstrap sample, it is known as to which data points are detected and which data points are censored. For each bootstrap sample, MLE was used to fit a parametric distribution. B estimates of the distribution parameters and of the fitted distributions were developed. Therefore, the sampling distribution of a given statistics are obtained based on the B estimates of the fitted distributions.[24]

2.4 Goodness-of-fit Test

The Kolmogorov-Smirnov (K-S) test is used to evaluate goodness-of-fit for data without censoring.[31, 37, 38] The fitted lognormal, gamma or Weibull distribution that has the smallest K-S test value was determined as the recommended parametric distribution. For censored data, a graphical comparison of the bootstrap confidence intervals for the cumulative distribution function (CDF) of the fitted distribution versus the empirical distribution of the data is used to evaluate goodness-of-fit. The larger the proportion of data contained within the confidence intervals, the more comfort an analyst will typically have regarding the goodness-of-fit. The distribution type whose bootstrap confidence intervals enclose the largest proportion of empirical data, especially for the upper percentiles that have the most influence on the mean estimate, is recommended as the preferred distribution.

2.5 Monte Carlo Simulation of Uncertainty in the Emission Inventory Model

The emission inventory for a pollutant is given by:

$$\text{EI} = \sum EF_i \times AF_i \qquad (1)$$

Where,

EF_i = emission factor from source i (mass emissions percent of activity)

AF_i = activity factor from source i (unit of activity)

Based upon selection of the preferred probability distribution model to represent inter-unit variability in the emission factor, uncertainty was estimated for the mean emission factor using bootstrap simulation. Uncertainty in the activity factor was estimated based upon judgment. The uncertainty in the emission inventory was simulated using Monte Carlo simulation[31, 38], resulting in an estimate of the probability distribution of uncertainty in the total inventory.

2.6 Sensitivity Analysis to Identify the Key Sources of Uncertainty

Sensitivity analysis based on rank correlation is used to identify the key sources of uncertainty. Rank correlation is a measure of the strength of the monotonic relationship between random variables.[31] Inputs that have a statistically significant correlation with the uncertainty in total inventory emissions were identified as sensitive inputs. The larger the magnitude of the correlation, the greater the sensitivity. Identification of the most highly sensitive inputs enables targeting of resources in future work to collect more or better information for specific source categories in order to reduce uncertainty in the total inventory.

3.0 Jacksonville Emission Inventory and Emission Factor Data

The Jacksonville point estimates of the total emissions for each of the six selected urban air toxics are given in Table 1. For each pollutant, between 11 and 16 source categories were identified as the priorities for uncertainty estimation, representing between 94 to 100 percent of total estimated emissions. Emission factor sample data upon which to base a statistical analysis were obtained for most but not all of these source categories. The specific emission source categories considered, the availability of data, and the results of the estimation of variability and uncertainty are summarized in Tables 2-7 for 1, 3-butadiene, mercury, arsenic, benzene,

formaldehyde and lead, respectively. The emission factor source categories are numbered separately for each pollutant based upon the highest to lowest contribution to total emissions.

Information sources for empirical sample data for emission factors include AP-42 background documents, EPA locating and estimating documents, the California Air Toxics Emission Factors data base and the Coordinating Research Council E55/E59 project report.[39-49] For some source categories, uncertainty was estimated based upon the product of two factors. For example, for onroad mobile sources, Frey and Zheng reported uncertainty in total hydrocarbon (THC) for gasoline engines and Bammi reported the uncertainty in the fraction of 1, 3-butadiene, benzene and formaldehyde in THC emissions.[13, 51] For nonroad mobile sources, Frey and Bammi reported the uncertainty in the THC emissions.[11, 14] For nonroad sources, the fraction of THC emitted as a given air toxic was estimated based upon the available data for onroad mobile sources.

For some source categories, directly relevant data were not available and surrogate data were used. For example, for 1, 3-butadiene, data for emission factors for wild fires are not available and uncertainty in the emission factors from forest fires were used as surrogates since both of the categories are large-scale natural combustion process that consumes various ages, sizes and types of outdoor vegetation[42]. For arsenic and lead, the uncertainty in the emission factors of external residual oil combustion sources is used as surrogate for that of external waste oil combustion sources considering the similarities between the fuel types. For benzene, emission factor sample data for gasoline use and architecture surface coating are not available. For the former, the uncertainty in a similar source category of bulk terminal gasoline solvent evaporation is used as surrogate. For the latter, assuming that the uncertainty in the fraction of VOC emitted as benzene is constant, the relative uncertainty in the VOC emission factors from architecture

surface coating is used as surrogate.[52] For formaldehyde, emission factor sample data for consumer/commercial solvent use and natural gas industrial combustion are not available. The uncertainty in VOC emission factors from consumer/commercial solvent use[52] is used as surrogate for the former and the uncertainty in formaldehyde emission factors from internal natural gas combustion engines is used as surrogate for the latter. The information sources for direct and surrogate data are cited for each source category in Tables 2-7 for the six pollutants.

As summarized in Table 1, the source categories for which directly relevant or surrogate data were available account for 80% to 99% of the total emission inventories.

4.0 Results

Inter-unit variability and uncertainty in the mean was quantified for each source category for which emission factor sample data or surrogates were available. Probabilistic emission inventories were developed considering the uncertainty in the emission factors and activity factors. The key sources of uncertainty were identified using sensitivity analysis.

4.1 Quantification of Variability and Uncertainty in Emission Factors

Each of the three candidate parametric distributions was fit to each emission factor data set, and a preferred distribution was selected per the methods previously discussed. The results for the uncensored case are described first, followed by results for the censored cases. The preferred distribution and their parameters for inter-unit variability, and the uncertainty in the mean emission factors are summarized in Tables 2-7 for the six pollutants for cases in which directly relevant data were available. Situations in which surrogate data were used are indicated. For some source categories, no sample data were available to support statistical analysis, but these comprise a small portion of the total inventory.

In most cases for uncensored data, the recommended distributions fit using MLE are not rejected by the K-S test. An example comparison of a fitted parametric distribution, its bootstrap confidence intervals, and the empirical data is given in Figure 1 for mercury emissions from pathological waste disposal (Case No. 6 in Table 3). The inter-unit variability is represented by a Weibull distribution and the data range over approximately five orders-of-magnitude. More than 95 percent of the data are enclosed by the 95 percent confidence interval and slightly more than half of the data are enclosed by the 50 percent confidence interval. Therefore, the Weibull distribution is deemed to be a good fit to the data.

For 1, 3 –butadiene Case Nos. 5 and 7 and benzene Case Nos. 3 and 10, MOMM was used instead of MLE to estimate the parameters since the former resulted in a smaller K-S test statistic.

For 1, 3-butadiene Case No. 11, no parametric distribution types were adequate fits and an empirical distribution was used instead in which each observed data point was assigned equal probability. For 1,3-butadiene Case No. 3, no single parametric distribution fit the data and a mixture lognormal distribution was used.[23] Figure 2 shows the bootstrap confidence intervals of the fitted mixture distribution in comparison to the sample data. Approximately 90 percent of the data are enclosed by the 95 percent confidence interval, with the largest significant discrepancies occurring in the lower portion of the distribution. This is a case in which a two component mixture substantially improves the accuracy of characterizing the upper tail of the distribution compared to a single component distribution. Although additional accuracy might be captured at the low end of the distribution with more components, the mean emission factor is more sensitive to the upper tail than to the lower tail. Therefore, refinements based upon more components would not lead to substantial improvements in the mean uncertainty estimate.

Moreover, with only 17 data points, the algorithms for fitting mixture distributions are likely to become unstable as more components are added, and the additional weights associated with more components will lead to addition uncertainty for portions of the fitted CDF.[23]

There were three uncensored cases, including Case Nos. 5 and 7 for 1, 3-butadience and Case No. 3 for benzene, for which the candidate distribution with the lowest K-S test statistic was rejected by the K-S test. For example, all these three cases have K-S test values about 0.32 compared to the critical K-S values equal to 0.30 at significance level of 0.05. However, a graphical comparison of the distribution with the smallest K-S test statistic similar to that of Figure 1 revealed adequate fits. For these three cases, all the empirical data were enclosed in the 95 percent confidence interval of the CDF and 3 out of 7 data points were enclosed in the 50 percent confidence interval.

For censored data sets, all the fits are good based on the graphical comparison of the fitted distribution and its confidence intervals to the data points. As an example, Figure 3 shows the variability and uncertainty in arsenic Case No. 1, external coal combustion boilers, estimated based upon a lognormal distribution. The data contain 26 detected and 3 censored values. Each censored value has a different detection limit; therefore, three detection limits are indicated. Twelve detected data are larger than the largest detection limit. They have exact ranks and exact cumulative probabilities. For detected data that are smaller in value than one or more detection limits, there is ambiguity in their respective ranks as well as cumulative probabilities. For example, an observed sample smaller than all three detection limits could have a rank as low as one or as large as four, depending upon the true but unknown values for each of the nondetected measurements. This ambiguity is depicted by a vertical solid line, instead of a point. The fitted distribution representing the inter-unit variability is shown as a white dashed line and it is

entirely enclosed in the 50 percent confidence interval. All the detected data points with exact cumulative probabilities and possible cumulative probabilities are enclosed in the 95 percent confidence interval.

The procedure of quantification of uncertainty in mobile source emission factors is different from other sources which have direct or surrogate data. Mobile sources contribute the most to 1, 3 – butadiene, benzene and formaldehyde emissions but are not reported as significant sources for arsenic and lead. For mercury, nonroad mobile source is the second largest source category but there is no available data for mercury emission factor from nonroad mobile source. For mobile source emission factors, previous work regarding uncertainty estimates has been done for THC for gasoline onroad sources as well as gasoline and diesel nonroad sources. The uncertainty in the fraction of 1, 3 –butadiene, benzene and formaldehyde in total organic gas emissions has also been estimated for onroad light duty gasoline vehicles (LDGV) and is used as the surrogate for other onroad and nonroad gasoline mobile sources. For diesel onroad mobile sources, the uncertainties in the THC emission factors and in 1, 3 – butadiene, benzene and formaldehyde fraction in THC emissions were quantified respectively based upon the empirical data with bootstrap simulation. The uncertainties in the toxics fractions were also used as the surrogates for nonroad diesel mobile sources. The variability and uncertainty in the THC emission factors and in 1, 3 – butadiene, benzene and formaldehyde fraction in THC emissions for diesel engines are given in Tables 2, 5 and 6 respectively. Based on available information, the uncertainty in the toxic emission factors were quantified based on the product of the uncertainty in THC emission factors and the uncertainty in the fraction of the toxics emissions in THC emissions. For gasoline mobile sources, only the uncertainty in the fraction of the toxics

emissions in TOG emissions is available; therefore, the uncertainty in the ratio of TOG to THC should be multiplied.

For benzene, Case 7 of bulk terminal gasoline solvent evaporation is an aggregation source category composed of subcategories for non-winter and winter time. For this case, the uncertainty in the emission factors of the aggregation source category was obtained by the weighted average uncertainties in the emission factors of the subcategories. Assuming winter time accounts for 3 months in a year, the weights assigned to non-winter and winter time are respectively equal to 0.75 and 0.25.

4.2 Development of Probabilistic Emission Inventories

Probabilistic emission inventories for the six pollutants in Jacksonville were developed based upon probabilistic mean emission factors and activity factors. The activity factor data of Jacksonville were not available. It is expected that there is uncertainty in the activity factors. However, in the absence of empirical data, a judgment was made to assign at least a minimal range of uncertainties to these activity factors. For each source category, a 95 percent confidence interval in the mean of activity factor was assumed ranging from minus 10 percent to plus 10 percent. Therefore, the normalized uncertainty estimates of the activity factors were generated from independent normal distributions with a mean of 1.0 and standard deviation of 0.05. The total uncertainty for each source category was calculated by multiplying the recommended uncertainty estimate of the emission factors by the uncertainty estimate of the corresponding activity factors. The resulting 95 percent confidence intervals in the emission inventory of each source category is given in Tables 2-7 for each pollutant, respectively. Probabilistic emission estimates were developed for source categories that have direct or surrogate data. The following equation was used to develop probabilistic emission inventory (PEI).

$$PEI = \sum_{i=1}^{n} [(UF_{EF,i})(UF_{AF,i})(EI_i)] \quad (1)$$

where,

PEI = Probabilistic emission inventory for sources which have direct or surrogate data

$UF_{EF, i}$ = Normalized uncertainty factor of emission factors for source i

$UF_{AF, i}$ = Normalized uncertainty factor of activity factors for source i

EI_i = Emission inventory from source i

In order to have insight on the influence of the correlation in the surrogates on the uncertainty results, two situations were assumed about the surrogates in the calculation: 100% correlated and uncorrelated. That is, the uncertainty factors of the emission factors using the same surrogates are randomized in the same order or different orders for different categories during Monte Carlo Simulation. Based on the source categories which have directly relevant or surrogate data, the quantified relative 95 percent confidence intervals were obtained for the six pollutants for the two situations. The results are shown in Table 8. From Table 8, the correlation between the surrogates does not have significant influence on the results. The largest uncertainty in the emission inventory occurs for arsenic with relative 95 percent confidence interval ranging from minus 83 percent to plus 243 percent, the smallest uncertainty in the emission inventory occurs for mercury with relative 95 percent confidence interval ranging from minus 25 percent to plus 30 percent.

4.3 Sensitivity Study of Identify Key Sources of Uncertainty

The results of sensitivity analysis are given in Tables 2-7 for each of the six pollutants, respectively, based upon the case study assuming that surrogates are correlated.

For 1, 3-butadiene, benzene and formaldehyde, the most sensitive source categories are onroad mobile sources. For mercury, arsenic and lead, the most sensitive source categories are external combustion boilers-electric generation-bituminous coal. These categories contribute the most to emissions and have the largest absolute uncertainty in the emission factors. The rank correlations for the most important categories range from 0.92 to 0.99 among the six pollutants, which indicate strong correlations.

Categories of secondary importance were identified. For formaldehyde, the second most important source category is aircraft with a rank correlation of 0.3. For lead, other significant source categories are external combustion with fuel oil or waste oil. Other source categories for each of the six pollutants typically have weak, such as less than 0.2, or statistically insignificant correlations. Therefore, the uncertainty in the emissions is primarily attributable to uncertainty in only one, two, or three source categories, depending upon the pollutant.

Since the calculation is based on the assumption that there is 100% correlation among the surrogates, the rank correlations of the source categories using the same surrogates are the same. For example, for lead, the uncertainty in the emission factors for waste oil external combustion uses that for fuel oil external combustion as surrogate and the two source categories have rank correlation of 0.244. Both of them are identified as key sources of uncertainty. Better data collection and reporting on fuel oil external combustion can reduce uncertainties in both source categories.

5.0 Discussion

In this paper, probabilistic emission inventories for six important urban air toxics were developed. The uncertainties in the emission inventories are quantified based upon available data. The source categories which have directly relevant or surrogate data account for more than

90% of the total emissions for 1, 3-butadiene, arsenic, benzene and formaldehyde, and account for more than 80% of the total emissions for mercury and lead. Therefore, the quantified uncertainty in the emissions which have data can be taken as a sufficient estimate of the uncertainty in the total emission inventory for all the pollutants.

Except for mercury, the 95 percent confidence intervals are on the order of minus 50 to plus 100 percent or more, indicating a factor of two or more for other 5 pollutants. The largest uncertainty occurs for arsenic with the 95 percent confidence interval ranging from minus 83 percent to plus 243 percent for arsenic. The large range of quantified uncertainty suggests that it is important to quantify uncertainty and that this portion of uncertainty should be taken into account when reporting and using emission factors.

Based on sensitivity study, for 1, 3-butadiene, mercury, arsenic and benzene, only one source category was identified as the major contributor to uncertainty in total emission inventory. For formaldehyde and lead, respectively two and four cases are identified as sensitive source categories. Onroad mobile sources are the dominate source for the three VOCs and external coal combustion sources are the dominate source for the three heavy metals. The reason is that these source categories are the largest source category and also have relative large uncertainty in the emission factors. Better data collection and reporting work should be prioritized for the key source categories.

The probabilistic emission inventory developed here could be improved in several ways pending availability of additional data or the incorporation of a more extensive expert elicitation component. For example, although biases in the mean emission factors are suspected, especially for fugitive emissions and as a result of process upset, insufficient data were available via which to quantify such biases. Other possible sources of bias include lack of representative data (e.g.,

measurements may have been for load or operating conditions not typical of annual average in-use activity) and the use of surrogate data for source categories in which data were lacking or not readily available. Expert elicitation could be used to encode judgments regarding the additional uncertainty associated with nonrepresentative or surrogate data. As new data become available, the assessment can be updated. A key obstacle to quantification of uncertainty based upon statistical data analysis is obtaining the necessary data. Often, data are measured and reported by multiple organizations. In the long term, the development of a protocol for archiving such data and making the data available would facilitate probabilistic analysis.

The uncertainty in the activity factors here is based on an approximate judgment to develop probabilistic emission inventory. In the long term, the quantifiable uncertainty in the activity factors should be incorporated when empirical data for activity factors are available for statistical analysis. For the cases lack of sample data, expert judgment may be required.

The results of this work demonstrate that random sampling error and measurement error are substantial sources of quantifiable uncertainty in the emission inventories of the example urban air toxics in Jacksonville area, especially for 1, 3-butadiene, arsenic, benzene, formaldehyde and lead. The positively skewed ranges of uncertainty appropriately account for the fact that emissions must be non-negative. The substantial ranges of uncertainty estimated here should be taken into account when conducting air quality modeling and exposure assessment. Furthermore, the identification of key sources of uncertainty in the inventory serves as an aid to prioritizing resources for additional data collection or research in order to reduce uncertainty.

6.0 Acknowledgements

The authors acknowledge the support of the Science to Achieve Results (STAR) grants program of the U.S. Environmental Protection Agency, which funded this work under Grant No. R826790. This paper has not been subject to any EPA review and therefore does not necessarily reflect the views of the Agency, and no official endorsement should be inferred.

7.0 References

1. National Research Council, *Rethinking the Ozone Problem in Urban and Regional Air Pollution*, National Academy Press: Washington, DC, 1991
2. National Research Council. *Science and Judgement in Risk Assessment*; National Academy Press: Washington, DC, 1994.
3. National Research Council. *Modeling Mobile-Source Emissions*; National Academy Press: Washington, DC, 2000.
4. IPCC, *Good Practice Guidance and Uncertainty Management in National Greenhouse Gas Inventories*, Intergovernmental Panel on Climate Change (IPCC), 2000. (available at http://www.ipcc nggip.iges.or.jp/public/gp/gpgaum.htm).
5. Frey, H.C., and D.S. Rhodes (1996), "Characterizing, Simulating, and Analyzing Variability and Uncertainty: An Illustration of Methods Using an Air Toxics Emissions Example," *Human and Ecological Risk Assessment: an International Journal*, 2(4):762-797 (December 1996).
6. Bennett, K., and R. Meyers (1997), "Uncertainty Analysis of AP-42 Emission Factors," *Proceedings, Emission Inventory: Planning for the Future*, Air & Waste Management Association, Pittsburgh, PA.

7. Pollack, A.K.; Bhave, P.; Heiken, J.; Lee, K.; Shepard, S.; Tran, C.; Yarwood, G.; Sawyer, R.F.; Joy, B.A. (1999), *Investigation of Emission Factors in the California EMFAC7G Model*; PB99-149718INZ, Prepared by ENVIRON International Corp, Novato, CA, for Coordinating Research Council: Atlanta, GA, 1999.
8. Anex, R. P., Lund, J. R. (1999). "Bayesian Estimation of Volatile Organic Emissions From Wood Furniture Coating." *J Environ. Eng.* 125, 92-96.
9. Anex, R. P., Lund, J. R, Grant, R. (1999). "A Maximum Entropy Approach to Estimating Emissions." *J. Air & Waste Manage. Assoc*. 49, 943-950.
10. El-Fadel, M., Massoud, M., Semerjian, L. (2001). Energy Related GHG Emissions: Assessment of Emission Factor Uncertainty. *World Resour. Rev*. 13, 61-73.
11. Frey, H.C., and S. Bammi (2002), "Quantification of Variability and Uncertainty in Lawn and Garden Equipment NO_x and Total Hydrocarbon Emission Factors," *Journal of the Air & Waste Management Association*, 52(4):435-448 (April 2002).
12. Frey, H.C., and J. Zheng (2002), "Quantification of Variability and Uncertainty in Utility NO_x Emission Inventories," *J. of Air & Waste Manage. Association*, 52(9):1083-1095 (September 2002).
13. Frey, H.C., and J. Zheng (2002), "Probabilistic Analysis of Driving Cycle-Based Highway Vehicle Emission Factors," *Environmental Science and Technology*, 36(23):5184-5191 (December 2002).
14. Frey, H.C., and S. Bammi (2003), "Probabilistic Nonroad Mobile Source Emission Factors," *ASCE Journal of Environmental Engineering*, 129(2):162-168 (February 2003).

15. Frey, H.C., and S. Li (2003), "Quantification of Variability and Uncertainty in AP-42 Emission Factors: Case Studies for Natural Gas-Fueled Engines," *Journal of the Air & Waste Management Association*, accepted for publication as of June 2003.
16. Abdel-Aziz, A., and H.C. Frey (2003), "Quantification of Hourly Variability in NO_x Emissions for Baseload Coal-Fired Power Plants," *Journal of the Air & Waste Management Association*, submitted January 2003, accepted July 2003.
17. Dickson, R.J., and Hobbs, A.D. (1989), "Evaluation of Emissions Inventory Uncertainty Estimation Procedures," Paper No. 89-24.8, Presented at the 82nd Annual Meeting of the Air & Waste Management Association, Anaheim, CA, June 25-30, 1995.
18. Guensler, R., D. Sperling and P.Jovanis (1991), "Uncertainty in the Emission Inventory for Heavy-Duty Diesel-Powered Trucks", Paper No. 91 -88.14, *Proceedings, 84th Annual Meeting*, Air and Waste Management Association, Vancouver, B. C. , June 1991.
19. Balentine, H.W. and R.J. Dickson (1995), "Development of Uncertainty Estimates for the Grand Canyon Visibility Transport Commission Emission inventory," Presented at the Air & Waste Management Association Conference on the Emission Inventory, Research Triangle Park, NC, October 1995.
20. Hanna, S.R., Z. Lu, H.C. Frey, N. Wheeler, J. Vukovich, S. Arunachalam, M. Fernau, and D.A. Hansen (2001), "Uncertainties in Predicted Ozone Concentrations due to Input Uncertainties for the UAM-V Photochemical Grid Model Applied to the July 1995 OTAG Domain," *Atmospheric Environment*, 35(5):891-903.
21. Frey, H.C., and D.S. Rhodes (1998), "Characterization and Simulation of Uncertain Frequency Distributions: Effects of Distribution Choice, Variability, Uncertainty, and

Parameter Dependence," *Human and Ecological Risk Assessment: an International Journal*, 4(2):423-468 (April 1998).

22. Frey, H.C., and D.E. Burmaster (1999), "Methods for Characterizing Variability and Uncertainty: Comparison of Bootstrap Simulation and Likelihood-Based Approaches," *Risk Analysis*, 19(1):109-130 (February 1999).
23. Zheng, J., and H.C. Frey (2003), "Quantification of Variability and Uncertainty Using Mixture Distributions: Evaluation of Sample Size, Mixing Weights and Separation between Components," *Risk Analysis*, submitted November 2002. Undergoing revision in response to comments.
24. Zhao, Y., and H.C. Frey (2003), "Quantification of Variability and Uncertainty for Censored Data Sets and Application to Air Toxic Emission Factors," *Risk Analysis*, submitted December 19, 2002. Undergoing revision in response to comments.
25. Smith, R.L., French, C.L., Murphy, D.L., and Thompson, R, *Ranking and Selection of Hazardous Air Pollutants for Listing under Section 112(k) of the Clean Air Act Amendments of 1990*, Technical Support Document of Integrated Urban Air Toxics Strategy, by EPA office of Air Quality Planning and Standards, 1999. http://www.epa.gov/ttn/atw/urban/urbanpg.html (accessed 10/1/ 2000).
26. US Environmental Protection Agency. *Guiding Principles for Monte Carlo Analysis*; EPA/630R-97/001; US Environmental Protection Agency: 1997.
27. US Environmental Protection Agency. *Guiding Principles for Monte Carlo Analysis*; EPA/630R-97/001; US Environmental Protection Agency: 1997.
28. EPA (1996c), "Summary Report for the Workshop on Monte Carlo Analysis," EPA/630/R-96/010, Risk Assessment Forum, Washington, DC.

29. Tilley, L. (2003), Personal communication with Lori Tilley, Regulatory and Environmental Services Department, City of Jacksonville, FL, via email, May 26th, 2003.

30. Rao, S.T., Ku, J.Y., and Rao, K. S. (1991), "Analysis of Toxic Air Contaminant Data Containing Concentrations Below the Limit of Detection," *Journal of the Air and Waste Management Association*, 41(4): 442-448.

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

32. Frey, H.C., Bharvirkar, R., and Zheng, J. (1999), "Quantitative Analysis of Variability and Uncertainty in Emissions Estimation," Prepared by North Carolina State University for Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC.

33. Seinfeld, J.H. (1986), *Atmospheric Chemistry and Physics of Air Pollution,* John Willey and Sons, New York.

34. Ott, W. (1990), "A Physical Explanation of the Lognormality of Pollutant Concentrations," *Journal of the Air and Waste Management Association*, 40: 1378-1383.

35. Ott, W. (1995), *Environmental Statistics and Data Analysis*, Lewis Publishers, Boca Raton, FL.

36. Efron, B. and Tibshirani, R.J. (1993), *An Introduction to Bootstrap,* Monographs on Statistics and Applied Probability, Chapman & Hall: New York.

37. Lu, HC., "Comparisons of Statistical Characteristic of Air Pollutants in Taiwan by Frequency Distribution," *Journal of the Air & Waste Management Association,* 53 (5): 608-616.

38. Morgan, M.G. and Henrion, M. (1990) *Uncertainty: A Guide to Dealing with Uncertainty in Qualitative Risk and Policy Analysis*, Cambridge University Press: New York, NY.

39. EPA, "Emission Factor Documentation for AP-42 Section 1.1 Bituminous and Subbituminous Coal Combustion," by Acurex Environmental Corporation, Edward Aul & Associates, Inc and E.H. Pechan and Associates, Inc, for Office of Air Quality Planning and Standards, Office of Air and Radiation, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1993. http://www.epa.gov/ttn/chief/ap42/ch01/bgdocs/b01s01.pdf

40. EPA, "Emission Factor Documentation for AP-42 Section 1.3 Fuel Oil Combustion," by Eastern Research Group, Office of Air and Radiation, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1993. http://www.epa.gov/ttn/chief/ap42/ch01/bgdocs/b01s03.pdf

41. EPA, "Emission Factor Documentation for AP-42 Section 2.6 Medical Waste Incineration," Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1993. http://www.epa.gov/ttn/chief/ap42/ch02/bgdocs/b02s03.pdf

42. EPA, "Locating and Estimating Air Emissions from Sources of 1, 3-butadiene," EPA-454/R-96-008, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1998. http://www.epa.gov/ttn/chief/le/butadn-a.pdf

43. EPA, "Locating and Estimating Air Emissions from Sources of Benzene," EPA-454/R-98-011, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1998. http://www.epa.gov/ttn/chief/le/benzene/index.html.

44. EPA, “Locating and Estimating Air Emissions from Sources of Formaldehyde,” EPA-450/4-91-012, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1991. http://www.epa.gov/ttn/chief/le/benzene/index.html

45. EPA, “Locating and Estimating Air Emissions from Sources of Mercury and Mercury Compounds,” EPA-454/R-97-012, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1997. http://www.epa.gov/ttn/chief/le/mercury1.pdf

46. EPA, “Locating and Estimating Air Emissions from Sources of Arsenic and Arsenic Compounds,” EPA-454/R-98-013, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1998. http://www.epa.gov/ttn/chief/le/arsnic_1.pdf

47. EPA, “Locating and Estimating Air Emissions from Sources of Lead and Lead Compounds,” EPA-454/R-98-006, Office of Air Quality Planning and Standards, U.S. Environmental Protection Agency, Research Triangle Park, NC, 1998. http://www.epa.gov/ttn/chief/le/lead1.pdf.

48. CARB, “California Air Toxics Emission Factor II Database,” Research Division, Air Resources Board, California Environmental Protection Agency, 2000. http://www.arb.ca.gov/emisinv/catef/catef.htm

49. CRC, “Heavy-duty Gasoline Vehicle Chassis Dynamometer Testing for Emissions Inventory,” Coordinating Research Council, 2003. http://www.crcao.com/

50. Frey, H.C. and Zheng, J. (2002), “Probabilistic Analysis of Driving Cycle-Based Highway Vehicle Emission Factors,” *Environmental Science and Technology*, 36(23): 5184-5161.

51. Bammi, S. (2001), *Quantitative Analysis of Variability and Uncertainty in On-Road and Non-Road Mobile Source Emission Factors,* Master thesis, Department of Civil Engineering, North Carolina State University, Raleigh, NC.
52. Li, S. (2002), *Development and Demonstration of a Methodology for Characterizing and Managing Uncertainty and Variability in Emission Inventories,* PhD dissertation, Department of Civil Engineering, North Carolina State University, Raleigh, NC.

Table 1. Summary of Available Data for 1, 3-butadiene, Mercury, Arsenic, Benzene, Formaldehyde and Lead Emission Inventories

Pollutant	Point Estimate of Emission Inventory	Percentage of Available Data Accounting for Total Emissions
1, 3-butadiene	173 tons/yr	99.1%
Mercury	500 lbs/yr	80%
Arsenic	1994 lbs/yr	98.5%
Benzene	764 tons/yr	97.4%
Formaldehyde	548 tons/yr	91.4%
Lead	2964 lbs/yr	83.5%

Table 2. Quantification of Variability and Uncertainty for 1, 3-butadiene Emission Inventory

Case No.	Emission Source Description	EI[a] (tons/yr)	Data Status[b]	Reference	n[c]	Variability in Emission Factor[d]	Uncertainty in Emission Factor (%)[e]	Uncertainty in Emission Inventory(%)[f]	Rank Correlation[g]
1	Onroad Gasoline Diesel b-1. THC b-2. 1, 3- butadiene/ THC	86.6 77.9 8.7	D D D D D	 31, 32 30	 24 24	 L (0.46, 1.46) W(1.66, 0.005)	(-81, 210) (-87, 235) (-61, 187) (-58, 151) (-23, 35)	(-82, 215)	**0.975**
2	Nonroad 2-stroke gasoline 4-stroke gasoline diesel	33.1 7.0 24.8 1.3	D/S D/S D/S D/S	 1, 32 1, 32 30, 33			(-29, 44) (-35, 45) (-35, 51) (-31, 39)	(-31, 44)	-0.08
3	Ships	30.6	D	23	17	ML: 0.82 (-1.09, 0.15) (-0.02, 0.20)	(-23, 27)	(-22, 28)	**0.108**
4	Aircraft	13.4	D/S	32			(-53, 76)	(-54, 79)	**0.168**
5	Prescribed Burning	6.45	D	23	7	G (2.26, 0.21)[h, i]	(-45, 57)	(-43, 52)	0.082
6	Trains	1.1							
7	Forest Fires	0.63	D	23	7	G (2.26, 0.21)[h, i]	(-41, 59)	(-44, 62)	-0.03
8	Wild Fires	0.63	S	Surrogate: Case 7			(-41, 59)	(-42, 58)	-0.03
9	Publicly Owned Treatment Works	0.483							
10	Vehicle Fires	0.0352							
11	Internal Combustion Engines	0.00062	D	20	16	Empirical	(-19, 16)	(-22, 19)	0.01

(Continued)

Table 2. Continued

[a.] Point estimate of emission inventory.

[b.] D = direct data; S = surrogate data; D/S = Direct data for THC, surrogate data for % of TOG (or THC) emitted as toxic.

[c.] Sample size, for censored data set, the number of non-detects is shown in parenthesis.

[d.] Inter-unit variability in emission factor. L: lognormal distribution; G: gamma distribution; W: Weibull distribution. The parameters of the distribution are given in parenthesis; ML: mixture lognormal distribution. 1st component = 0.82, parameters for the first component are (-1.09, 0.15), 2nd component = 1-0.82 = 0.18, parameters for the second component are (-0.02, 0.20).

[e.] The 95 % confidence interval relative to the mean is given. For the cases in which there is no information of variability in emission factors while there is information of uncertainty in the emission factor, the latter is based upon surrogate or previous work.

[f.] The 95 % confidence interval relative to the mean is given.

[g.] Rank correlation between the uncertainty in the total emissions and the uncertainty in the emission factors for each source category with correlated surrogates. Statistically significant correlations are shown in boldface. Correlation of less than 0.088 in magnitude are not statistically significant based upon the inverse Fisher transformation.

[h.] Rejected by K-S test, but the fit is judged to be adequate.

[i.] Fit parametric distribution with MOMM method instead of MLE since MOMM results in a smaller K-S test.

Table 3. Quantification of Variability and Uncertainty for Mercury Emission Inventory*

Case No.	Emission Source Description	EI[a] (lbs/yr)	Data Status[b]	Reference	n[c]	Variability in Emission Factor[d]	Uncertainty in Emission Factor (%)[e]	Uncertainty in Emission Inventory(%)[f]	Rank Correlation[g]
1	External Combustion Boilers - Electric Generation - Bituminous Coal	352.62	D	20	29 (3)	W (1.17, 8.56)	(-28, 34)	(-28, 34)	**0.934**
2	Nonroad	40							
3	Internal Combustion Engines - Electric Generation/Industrial - Natural Gas	33.56							
4	Waste Disposal - Solid Waste Disposal - Government - Other Incineration	18.93	D	26	5	W (1.15, 0.79)	(-62, 88)	(-60, 91)	0.034
5	External Combustion Boilers - Electric Generation/Industrial/ Commercial - Residual Oil	16.07	D	21	13 (11)	L (1.66, 1.07)	(-31, 32)	(-35, 30)	0.100
6	Waste Disposal - Solid waste disposal - Commercial/Institutional – Pathological	10.67	D	22	40	W (0.40, 14.5)	(-67, 136)	(-68, 142)	0.073
7	Human Crematory	8.32							
8	Distillate Oil – Industrial	5.52							

* The footnotes are the same as Table 2

Table 4. Quantification of Variability and Uncertainty for Arsenic Emission Inventory*

Case No.	Emission Source Description	EI[a] (lbs/yr)	Data Status[b]	Reference	n[c]	Variability in Emission Factor[d]	Uncertainty in Emission Factor (%)[e]	Uncertainty in Emission Inventory(%)[f]	Rank Correlation[g]
1	External Combustion Boilers - Electric Generation - Bituminous Coal	1741.87	D	20	29 (3)	L (-1.28, 2.35)	(-91, 246)	(-91, 272)	**0.992**
2	External Combustion Boilers - Electric Generation/Industrial/ Commercial - Residual Oil	187.66	D	21	13 (3)	W (1.03, 1.07)	(-47, 59)	(-48, 56)	**0.187**
3	External Combustion Boilers - Electric Generation - Waste Oil	22.8	S	Surrogate: Case 2			(-47, 59)	(-48, 63)	**0.187**
4	Residual Oil – Industrial	10.32	S	Surrogate: Case 2			(-47, 59)	(-46, 59)	**0.187**
5	Internal Combustion Engines - Electric Generation - Distillate Oil	7.58							
6	Distillate Oil – Industrial	7.36							
7	Ships	4.9							
8	External Combustion Boilers - Electric Generation/Industrial/ Commercial/Space Heaters - Natural Gas	3.21							

(Continued)

Table 4. Continued

Case No.	Emission Source Description	EI[a] (lbs/yr)	Data Status[b]	Reference	n[c]	Variability in Emission Factor[d]	Uncertainty in Emission Factor (%)[e]	Uncertainty in Emission Inventory(%)[f]	Rank Correlation[g]
9	External Combustion Boilers - Electric Generation/Industrial/ Commercial - Distillate Oil	3.10							
10	Distillate Oil – Commercial	2.08							
11	Natural Gas – Industrial	1.40							
12	Industrial Processes - Minerals Production - Asphalt Concrete	0.51	D	27	4	L (1.62, 0.57)	(-44, 73)	(-46, 72)	-0.02

* The footnotes are the same as Table 2

Table 5. Quantification of Variability and Uncertainty for Benzene Emission Inventory*

Case No.	Emission Source Description	EI[a] (tons/yr)	Data Status[b]	Reference	n[c]	Variability in Emission Factor[d]	Uncertainty in Emission Factor (%)[e]	Uncertainty in Emission Inventory(%)[f]	Rank Correlation[g]
1	Onroad Gasoline Diesel b-1. THC b-2. 1, 3- butadiene/ THC	435.0 391.5 43.5	D D D D D	31, 32 30	24 24	L (0.46, 1.46) W (3.36, 0.01)	(-81, 218) (-87, 237) (-59, 166) (-58, 151) (-13, 13)	(-88, 249)	**0.984**
2	Nonroad 2-stroke gasoline 4-stroke gasoline diesel	247.68 84.21 148.61 14.86	D/S D/S D/S D/S	1, 32 1, 32 30, 33			(-25, 33) (-32, 40) (-34, 46) (-26, 30)	(-25, 37)	**0.166**
3	Prescribed burning	17.93	D	24	7	G (2.21, 0.59)[h, i]	(-45, 53)	(-46, 56)	0.034
4	Gasoline use	17.27	S	Surrogate: Case 7			(-62, 155)	(-62, 156)	-0.026
5	Aircraft	15.10	D/S	32				(-50, 79)	0.088
6	Surface cleaning/decreasing	15.06							
7	Petroleum and solvent evaporation-bulk terminal gasoline a. non winer b. winter	4.51	D D D	24 24 24	 11 11	 L (-3.86,1.44) L (-3.53,1.43)	(-62, 155)	(-62, 151)	-0.026
8	Surface coating-auto refinishing	3.29							
9	External combustion boilers – electronic generation – bituminous coal	2.76	D	20	18 (1)	L (-2.38, 2.36)	(-93, 411)	(-93, 422)	0.020

(Continued)

Table 5. Continued

Case No.	Emission Source Description	EI[a] (lbs/yr)	Data Status[b]	Reference	n[c]	Variability in Emission Factor[d]	Uncertainty in Emission Factor (%)[e]	Uncertainty in Emission Inventory(%)[f]	Rank Correlation[g]
10	Wildfires	1.76	D	24	6	W (1.54, 1.67)[i]	(-47, 54)	(-47, 58)	0.086
11	Surface coating - architecture	1.60	S	Surrogate: VOC			(-8.4, 8.6)	(-13, 13)	-0.08
12	Vehicle fires	0.47							
13	Petroleum and solvent evaporation – storage – fuels other than gasoline	0.45							
14	Publicly Owned Treatment Work	0.32							
15	Industrial processes – minerals production – asphalt concrete	0.22	D	24	5	G (1.14, 1.74)	(-60, 90)	(-62, 93)	-0.04
16	Petroleum and solvent evaporation – transportation and marketing of petroleum products – marine vessels	0.13	D	24	9	L (-4.17, 0.54)	(-32, 40)	(-33, 39)	0.07

* The footnotes are the same as Table 2

Table 6. Quantification of Variability and Uncertainty for Formaldehyde Emission Inventory*

Case No.	Emission Source Description	EI[a] (tons/yr)	Data Status[b]	Reference	n[c]	Variability in Emission Factor[d]	Uncertainty in Emission Factor (%)[e]	Uncertainty in Emission Inventory(%)[f]	Rank Correlation[g]
1	Onroad Gasoline Diesel b-1. THC b-2. 1, 3- butadiene/ THC	216.0 194.4 21.6	D D D D D	31, 32 30	24 24	L (0.46, 1.46) W (1.43, 0.09)	(-80, 210) (-87, 224) (-63, 166) (-58, 151) (-28, 30)	(-82, 217)	**0.919**
2	Nonroad 2-stroke gasoline 4-stroke gasoline diesel	167.8 8.4 45.3 114.1	D/S D/S D/S D/S	1, 32 1, 32 30, 33			(-26, 36) (-36, 51) (-39, 59) (-34, 42)	(-27, 33)	**0.090**
3	Aircraft	111.0	D/S	32			(-53, 80)	(-53, 84)	**0.328**
4	Prescribed burning	41.04							
5	Wild fires	4.02							
6	External combustion boilers – electric generation/industrial/commercial-residual oil	2.34	D	21	13 (8)	W (0.30, 3.29)	(-95, 164)	(-95, 169)	0.003
7	Internal combustion engines – electric generation/industrial-natural gas-turbines[c]	1.05	D	29	7	W (1.41, 2.01)[i]	(-50, 60)	(-49, 58)	-0.02
8	Industrial processes – minerals production – asphalt concrete[c]	0.94	D	29	6	G (8.53, 0.33)	(-25, 29)	(-27, 30)	-0.02
9	Structure fires	0.70							

(Continued)

Table 6. Continued

Case No.	Emission Source Description	EI[a] (lbs/yr)	Data Status[b]	Reference	n[c]	Variability in Emission Factor[d]	Uncertainty in Emission Factor (%)[e]	Uncertainty in Emission Inventory(%)[f]	Rank Correlation[g]
10	External combustion boilers – electric generation/industrial/ commercial/space heaters – natural gas	0.60							
11	Internal combustion engines – engines testing – aircraft engine testing	0.58							
12	External combustion boilers – electric generation – bituminous coal	0.51	D	20	14 (5)	L (-0.59, 1.44)	(-77, 209)	(-78, 211)	0.018
13	Consumer/commercial solvent use	0.49	S	Surrogate: VOC			(-7.8, 8.5)	(-17, 18)	0.06
14	Natural gas – industrial	0.26	S	Surrogate: Case 7			(-50, 60)	(-50, 61)	-0.02
15	External combustion boilers – electric generation/industrial/ commercial – distillate oil	0.17	D	29	18	L (-1.65, 1.20)	(-50, 96)	(-50, 109)	0.07
16	Residual oil - industrial	0.13	D	21	13 (8)	W (0.30, 3.29)	(-95, 164)	(-95, 172)	0.003

* The footnotes are the same as Table 2

Table 7. Quantification of Variability and Uncertainty for Lead Emission Inventory*

Case No.	Emission Source Description	EI[a] (lbs/yr)	Data Status[b]	Reference	n[c]	Variability in Emission Factor[d]	Uncertainty in Emission Factor (%)[e]	Uncertainty in Emission Inventory(%)[f]	Rank Correlation[g]
1	External combustion boilers – electric generation – bituminous coal	1780	D	20	27	L (-0.065, 1.82)	(-69, 258)	(-69, 249)	**0.956**
2	External combustion boilers – electric generation – waste oil	457	S	Surrogate: Case 4			(-52, 53)	(-52, 56)	**0.244**
3	Miscellaneous organics	260							
4	External combustion boilers – electric generation/industrial/ commercial – residual oil	215	D	21	13 (4)	W (0.85, 1.28)	(-52, 53)	(-52, 55)	**0.244**
5	Industrial processes-minerals production - glass manufacture	103							
6	Internal combustion engines – electric generation – distillate oil	89.7							
7	Distillate oil – industrial	16.54							
8	Residual oil – industrial	11.82	S	Surrogate: Case 4			(-52, 53)	(-52, 54)	**0.244**

(Continued)

Table 7. Continued

Case No.	Emission Source Description	EI[a] (lbs/yr)	Data Status[b]	Reference	n[c]	Variability in Emission Factor[d]	Uncertainty in Emission Factor (%)[e]	Uncertainty in Emission Inventory(%)[f]	Rank Correlation[g]
9	External combustion boilers – electric generation/industrial/ commercial – distillate oil	6.97							
10	Ships	5.74							
11	Waste disposal – solid waste disposal – commercial/institutional – pathological	4.90	D	22	48	L (1.62, 1.02)	(-29, 45)	(-30, 46)	-0.038
12	Distillate oil – commercial	4.68							
13	Natural gas – industrial	1.89							
14	Industrial processes – minerals production – asphalt concrete	1.43	D	28	5	W (1.35, 1.54)	(-52, 78)	(-53, 79)	-0.040

* The footnotes are the same as Table 2

Table 8. Results of the Uncertainties in The Total Emission Inventories for 1, 3-butadiene, Mercury, Arsenic, Benzene, Formaldehyde and Lead

Pollutant	95 Percent Confidence Interval in the Emission Inventories (%)	
	Correlated Surrogates	Uncorrelated Surrogates
1, 3-butadiene	(-46, 108)	(-46, 108)
Mercury	(-25, 30)	(-25, 30)
Arsenic	(-83, 243)	(-83, 243)
Benzene	(-56, 146)	(-54, 141)
Formaldehyde	(-42, 89)	(-42, 89)
Lead	(-54, 175)	(-52, 177)

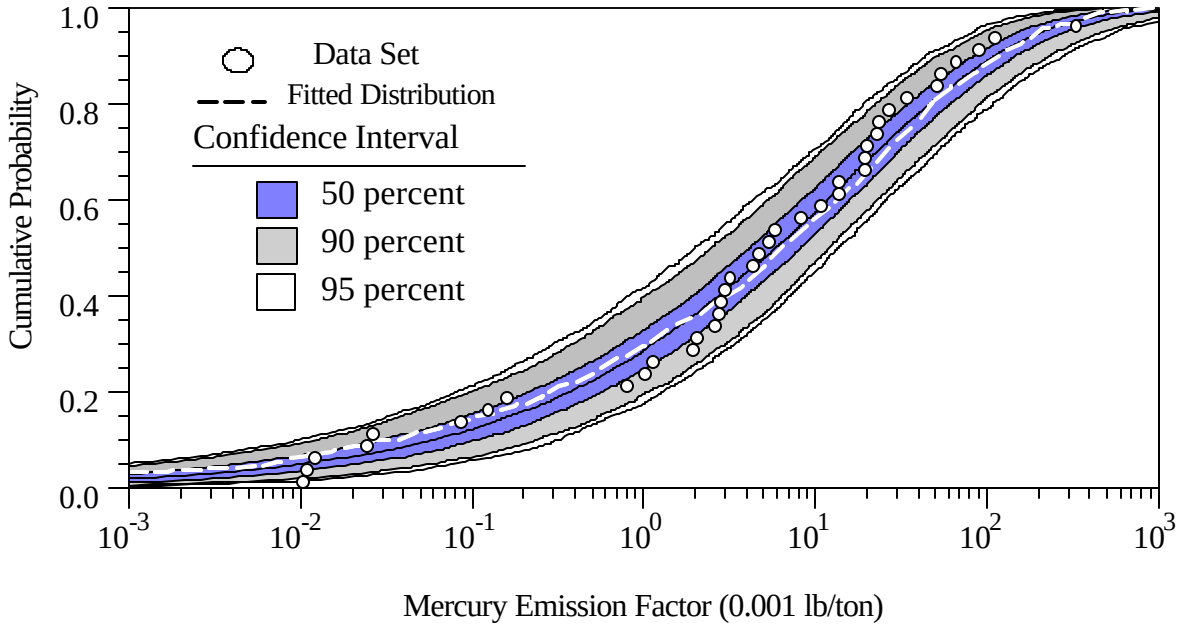


Figure 1. Variability and Uncertainty in Mercury Emission Factor from Pathological Waste Disposal Estimated Based Upon a Weibull Distribution

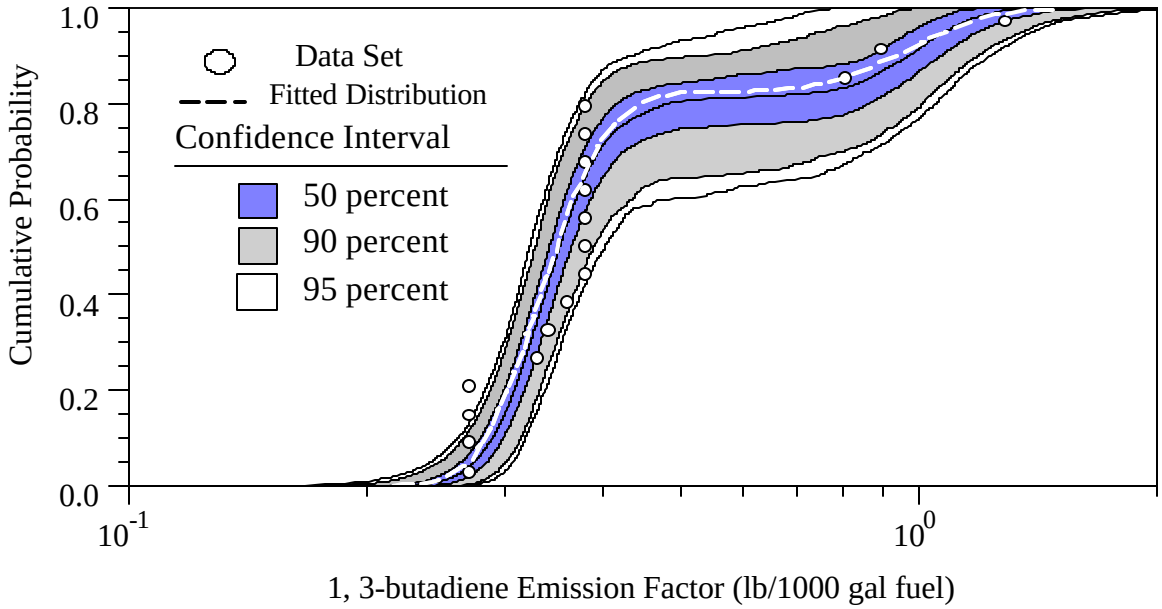


Figure 2. Uncertainty in 1, 3-butadiene Emission Factor from Ships Estimated Based Upon Mixture Lognormal Distribution

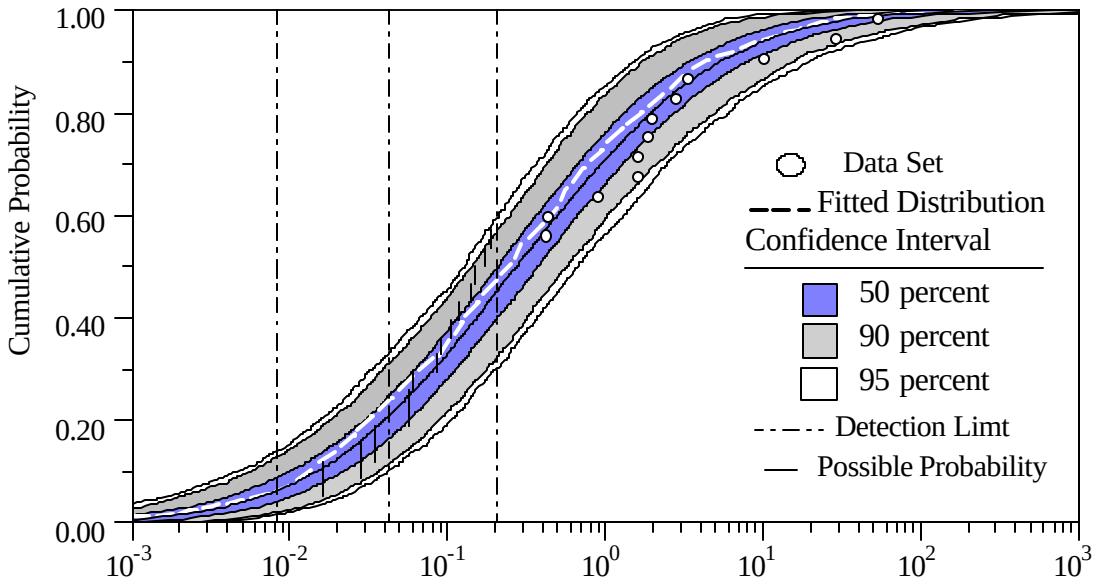


Figure 3. Variability and Uncertainty in Arsenic Emission Factor from Coal Combustion Estimated Based Upon a Lognormal Distribution

Part VI

CONCLUSIONS AND RECOMMENDATIONS

1.0 Conclusions and Recommendations

In this section, the conclusions and implications regarding the development and application of the MLE/Bootstrap method are first drawn and the motivating questions regarding censored data issues are answered. Second, the conclusions and implications regarding the development of probabilistic urban air toxics emission inventories are drawn and the motivating questions are answered. Finally, the recommendations for future work are presented.

1.1 Conclusions Regarding MLE/Bootstrap Method for Censored Data

The MLE method is a flexible and robust method for fitting parametric distributions to censored data. The fitted parametric distribution represents the inter-unit variability in censored data. It provides asymptotically unbiased estimate of statistics, such as the mean for censored data. It has been demonstrated by application to different distribution types, including the lognormal, gamma and Weibull distributions, and to data with different statistical characteristics, such as sample sizes from 20 to 100, coefficients of variation from 0.5 to 2, and differing amounts and types of censoring, including single and multiple detection limits involving as much as 60 percent of the distribution.

With using "bootstrap pairs" in empirical bootstrap simulation, the uncertainty in the estimated CDF and statistics for censored data can be quantified. The capability of bootstrap simulation to estimate uncertainty in statistics of censored data sets represented by MLE fitted distributions was demonstrated for the cumulative distribution function, mean, median, and 90^{th} percentile. The MLE/Bootstrap method for quantification of censored data takes into account uncertainty associated with random sampling error, measurement error and censoring issues.

The same as for uncensored data, the sample size and inter-unit variability influence the estimated uncertainty for censored data. The smaller the sample size and the larger the variability

in the data, the larger the uncertainty is in the mean and in the CDF for censored data. However, the uncertainty associated with censoring is more complicated for different situations. In Part II, the factors that influence uncertainty results for censored data with one, two and three detection limits are evaluated. It is found that the statistics that are sensitive to large values of a data set, such as the mean, may not be particularly sensitive to uncertainty associated with left-censoring. It is clear that the ranges of uncertainty of the portions of the CDF below detection limit become larger with more censoring. In Part III, insights regarding uncertainty factors for censored data are obtained based on the comparison of the uncertainty in the mean for the censored data and modified data for 16 different cases with single or multiple detection limits. When the uncertainty associated with censoring becomes considerable, it maybe worthwhile to consider approaches for reducing the detection limit in future data collection such as by increasing the sample volume or by using more sensitive instruments. The censoring issues become critical in the following situations:

- When there are some detection limits that are larger than the largest detected value;
- When there is a large amount of censoring;
- When the data set are not very positively skewed and thus the influence by censoring on the mean is more considerable;
- When statistics not sensitive to large values are of interest, such as the median and the uncertainty in the median.

In Part III, the MLE/bootstrap method is applied to 16 cases of urban air toxics emission factors with different sample sizes, degree of censoring and inter-unit variability. The MLE/Bootstrap provides consistent results for censored data with single or multiple detection

limits. Even for data with censoring degrees as high as 80%, the MLE/bootstrap method gives reliable results.

The MLE/bootstrap method can be applied to left-censored data sets in other fields except urban air toxics data. For example, the MLE/bootstrap method can be used to analyze the left-censored pyrene concentration data mentioned in Section 1.3. Compared to the non-parametric Kaplan-Meier estimator, the MLE/bootstrap method is more reliable for small sample sizes and for the results in the tail of the distributions.

The MLE/bootstrap method can be applied to the bioassay data for which left censoring is a characteristic. It can provide the best estimated statistics of interest as well as the uncertainty result in the estimated statistics for the bioassay data.

Furthermore, there are large amount of right censored data in survival analysis. With a simple transformation, right-censored data can be transformed to left-censored data, thus, the MLE/bootstrap method can be applied to right-censored data as well. For example, right censored survival data can be transformed to left censored data if each detected and censored point in the data set is subtracted by a large value. The MLE method for interval-censored data is also available. For example, Equation 8 in Part II is the log-likelihood function for interval-censored data. After changing the code of log-likelihood function to interval-censoring, the MLE/bootstrap method can also be applied to interval-censored data as well.

1.2 Conclusions Regarding the Development of Probabilistic Urban Air Toxics Emission Inventories for Houston, TX and Jacksonville, FL

Probabilistic emission inventories were developed for benzene, formaldehyde, 1,3-butadiene, and chromium based on the 1996 Houston inventory as well as for 1, 3-butadiene, mercury, arsenic, benzene, formaldehyde and lead based on the 2000 Jacksonville inventory.

The uncertainty in the mean emission factor for different source categories and different pollutants ranges from as small as approximately plus or minus 2 percent to as large as minus 99 to plus 600 percent. Typically, there is a factor of two uncertainty in the mean emission factors for urban air toxics. The relative uncertainty in the total emission inventories range from as small as approximately minus 20 to plus 34 percent, as in the case of chromium for Houston, to as large as minus 83 to plus 243 percent, as in the case of arsenic for Jacksonville. The typical large uncertainty in the urban air toxics emission factors and emission inventories indicate it is important to quantify uncertainty when collecting urban air toxics emission factor and inventory data. Efforts of better data collection and reporting work to reduce uncertainty should be prioritized on the identified key sources.

The quantified ranges of uncertainty for urban air toxics emissions take into account random sampling error and measurement error in emission factors. The former is influenced by the sample size and inter-unit variability for each emission source category. Random measurement error is accounted for because the observed variability in the data includes both the true variability and the random component of measurement error, which in turn influences the range of the sampling distribution of the mean.

Surrogate data were used to develop probabilistic urban emission inventories when directly relevant data were not available. In fact, in this procedure, uncertainty caused by using surrogate is introduced. Now there is not enough information to quantify uncertainty caused by using surrogate data based upon statistical analysis of empirical data. Since the source categories for the surrogate data and target data are all reasonably similar to each other in all the cases, such as utility boilers versus industrial boilers, nonroad mobile source versus onroad mobile source, and the correlation among the surrogates does not influence the uncertainty results significantly,

it is reasonable to assume that the uncertainty caused by using surrogate data does not contribute much to the uncertainty in the total emission inventories. It is recommended that the uncertainty caused by using surrogate data should be quantified by expert judgment and taken into account in the quantified uncertainty in the total emission inventories in the future.

Although the distributions of the source categories for the two cities are different for a given pollutant, the uncertainty results for a given pollutant were comparable to each other. For example, with correlated surrogates, the uncertainty in benzene emission ranges from minus 46 percent to plus 108 percent from the mean value in Houston area and ranges from minus 56 percent to plus 146 percent from the mean value in Jacksonville. The uncertainty in formaldehyde emission inventory ranges from minus 36 percent to plus 69 percent in Houston and ranges from minus 42 percent to plus 89 percent for Jacksonville. The key sources of uncertainty for benzene and formaldehyde emission inventories for both of the cities are identified as onroad mobile source. For arsenic, the uncertainty in the emission inventories ranges from minus 75 percent to plus 224 percent for Houston and ranges from minus 83 percent to plus 243 percent for Jacksonville. The key source of uncertainty for arsenic emission inventories for both of the cities is identified as external combustion boilers. This indicates that the uncertainties in the total emission inventories for a given pollutant mainly result from the key source categories and are not significantly influenced by other source categories.

1.3 Recommendations for Future Work

Based on the research of this dissertation, key recommendations for future work are:

(1) MLE/Bootstrap has been proved as a statistically rigorous, robust and asymptotically unbiased method that can be used to make inferences for a wide variety of situations, including different types of distributions, coefficients of variation, sample sizes, and amounts of censoring

with either single or multiple detection limits. Compared to simplified conventional methods, which are biased, the MLE/bootstrap method is more computationally intensive. However, unlike conventional methods, the MLE/bootstrap enables estimation of uncertainty for any statistic, including the influence of uncertainty associated with censoring itself. The MLE/bootstrap method can be incorporated as part of a two dimensional framework in which variability and uncertainty are distinguished. The MLE/Bootstrap method is therefore recommended to estimate the mean for censored urban air toxic emission factor data and to quantify the variability and uncertainty in censored urban air toxic emission factors.

(2) The wide range of quantified uncertainty in urban air toxic emission factors indicates the importance of quantifying this uncertainty. Therefore, information on which statistical analysis is based needs to be tested and well documented for urban air toxic emissions factors. For example, petroleum refineries-catalytic cracking is the largest source category for chromium emission inventory in Houston, but the quantified uncertainty in this category is suspiciously small based on the available data. For this category, future work needs to be done when new data becomes available.

(3) In the procedure of development of a probabilistic urban air toxic emission inventory, the uncertainty in the activity factors estimated here is based on an approximate judgment. In the long term, the quantifiable uncertainties in the activity factors should be incorporated based upon expert judgment.

(4) The probabilistic emission inventory developed here could be improved in several ways pending availability of additional data or the incorporation of a more extensive expert elicitation component. For example, although biases in the mean emission factors are suspected, especially for fugitive emissions and as a result of process upset, insufficient data were available

via which to quantify such biases. Other possible sources of bias include lack of representative data and the use of surrogate data for source categories in which data were lacking or not readily available. Expert elicitation is recommended to be used to encode judgments regarding the additional uncertainty associated with nonrepresentative or surrogate data. As new data become available, the assessment can be updated.

(5) A key obstacle to quantification of uncertainty based upon statistical data analysis is obtaining the necessary data. Often, data are measured and reported by multiple organizations. In the long term, the development of a protocol for archiving such data and making the data available is recommended to facilitate probabilistic analysis.

(6) In this project, the uncertainty in the total emission inventory is quantified based upon the source categories which have directly relevant or surrogate data. However, since the uncertainty in the emission inventory is caused by only a few source categories, it is actually not necessary to quantify the uncertainty in the emission inventories for all the source categories. For example, for all of the ten probabilistic emission inventories which have been developed, the key sources of uncertainty are among the largest five source categories. The other source categories either have insignificant correlation or have weak correlation to the uncertainty in the total emission inventories even if the uncertainty in the emission inventory from that source category is large. Thus, if just considering the largest five source categories to quantify the uncertainty in the total emission inventories, the results of the uncertainty in the total emission inventory will not be significantly different from that when considering all the source categories. Furthermore, based on the ten available emission inventories, it is found that the amounts of the urban air toxics emissions are quite different from source to source and the sum of the emissions from the largest five source categories contributes more than 78% to the total emissions even if there are

27 major source categories in the emission inventory. Thus if the emissions from the several largest source categories are considerably larger than the others, it is possible to quantify the uncertainty in the total emission inventory just based on the largest several source categories. Therefore, the amount of work will be reduced.

(7) For some source categories, such as for mobile sources, some urban air toxics (e.g. heavy metals) emissions are very low, all the data points may be censored. The available optimization program for MLE method can not provide stable results for this situation. Further study regarding statistical and engineering method to deal with censored data with 100% censoring needs to be carried on.

In concise summary, the innovative MLE/bootstrap method is a robust method to deal with left, right and interval censored data even if the censoring degree is as high as 80% and the sample size is as small as eight. Since censored data is a common problem in urban air toxic emission inventories, the MLE/bootstrap method is recommended to analyze censored urban air toxics data and to develop probabilistic urban air toxics emission inventories.