Methods for Quantifying Variability and Uncertainty in AP-42 Emission Factors: Case Studies for Natural Gas-Fueled Engines

H. Christopher Frey, Song Li
Department of Civil Engineering, North Carolina State University, Raleigh, NC 27695
frey@eos.ncsu.edu, sli3@eos.ncsu.edu

Cite as: Frey, H.C., and S. Li, "Methods for Quantifying Variability and Uncertainty in AP-42 Emission Factors: Case Studies for Natural Gas-Fueled Engines," *Proceedings, U.S. Environmental Protection Agency Emission Inventory Conference*, Atlanta, GA, April 2002

ABSTRACT

Quantitative methods for characterizing variability and uncertainty were applied to case studies of emission factors for stationary natural gas-fueled internal combustion engines. NO_x and Total Organic Carbon (TOC) emission data sets for lean burn engines were analyzed. Data were available for uncontrolled engines and for engines with pre-combustion chamber (PCC) and "clean burn" NO_x control approaches. For each data set, parametric probability distributions were fit to the data to represent inter-engine variability in emissions. Bootstrap simulation was used to quantify uncertainty in the fitted distribution and in the mean emission factor. Some methodological challenges were encountered in analyzing the data. For example, in one instance, only five data points were available, with each data point representing a different market share. Therefore, an approach was developed in which a parametric distribution was fitted to population-weighted data. The range of uncertainty in mean emission factors ranges from as little as approximately plus or minus 10 percent to as much as minus 60 percent to plus 80 percent, depending on the pollutant, control technology, and nature of the available data. The wide range of uncertainty in some emission factors emphasizes the importance of recognizing an accounting for uncertainty in emissions estimates. The skewness in some uncertainty estimates illustrates the importance of using a numerical simulation approach that does not impose restrictive symmetry assumptions on the confidence interval for the mean. In this paper, the probabilistic analysis method, the data sets, the results of the analyses, and key findings and recommendations are presented.

INTRODUCTION

The use of quantitative methods for characterizing variability and uncertainty applied to emission factors is demonstrated here. Emission factors are a key input to emission inventories. Emission inventories, in turn, are widely used for regulatory and air quality management purposes. However, the uncertainty in emission factors, and in emission inventories, is typically not known. Therefore, it is not known, in many cases, how robust regulatory or management decisions are with respect to uncertainty. If management decisions are based upon point estimates of emissions that are biased, or if the range of uncertainty in emissions is much larger than any predicted change in emissions resulting from an air quality management strategy, then the decision-making process for developing management strategies will be ineffective. This paper focuses on one of the fundamental starting points for characterizing uncertainty in emission inventories, which is the emission factor. The case study application is stationary natural gas-fueled reciprocating engines.

Variability and Uncertainty

Emissions vary from one specific source to another (e.g., one engine to another) and for a given source because of variations in design, feedstock compositions, ambient conditions, and other operating conditions. Thus, there is typically some inherent variation in emissions that is revealed by measurements on multiple specific emission sources or by repeated measurements of the same emission source.

Uncertainty refers to lack of knowledge regarding the true but unknown value of a quantity, such as the true but unknown population average emission factor for a particular source category.[1,2] The average emission factor is subject to uncertainty for several possible reasons: (1) random sampling error; (2) measurement errors; (3) non-representativeness of available data; and/or (4) lack of information. There is also the possibility that there are data entry mistakes. In this paper, the main focus is on quantification of random sampling error, which is the statistical random fluctuation in any statistic estimated from a finite random sample of data. Any statistic estimated from a random sample of data, such as the mean, is itself a random variable. The probability distribution for a statistic is referred to as the sampling distribution. The sampling distribution can be used to develop confidence intervals for a statistic. In this paper, sampling distributions are used as a method for quantifying uncertainty associated with random sampling error.

Estimation of Uncertainty in Emission Factors

Current practice in emission inventory work is typically to ignore uncertainty. As a surrogate for uncertainty estimates, some emission factors are accompanied by data quality ratings.[3] A method for qualitatively rating emission inventories, known as the Data Attribute Rating System (DARS) has been developed by EPA.[4] Some sources of uncertainty are difficult to quantify, such as non-representativeness of a data set. Therefore, there will always be a role for qualitative statements regarding non-quantifiable sources of uncertainty. However, qualitative rating systems should be used in combination with quantitative approaches.

There is growing recognition of the importance of quantitative uncertainty analysis in environmental modeling and assessment. For example, the U.S. EPA has developed guidelines for Monte Carlo analysis of uncertainty.[5] The National Academy of Sciences has repeatedly recommended to EPA that quantitative analysis of uncertainty be included in a variety of applications.[6,7] The Intergovernmental Panel on Climate Change (IPCC) has recently issued good practice guidance for quantifying uncertainty in emission inventories.[8]

As part of previous and ongoing work, research is underway to develop and demonstrate improved methods for quantifying uncertainty in emission inventories. In the area of mobile source emissions, for example, Kini and Frey developed quantitative estimates of uncertainty associated with Mobile5b emission factor model estimates of light duty gasoline vehicle base emissions and speed-corrected emissions.[9] Pollack *et al*. performed a similar study on California's EMFAC7G highway vehicle emission factor model.[10] Frey *et al*. revisited the earlier analysis of Mobile5b emission factor estimates to include uncertainties associated with temperature corrections.[11] Frey and Bammi estimated uncertainty in the emission factors for a non-road source category of lawn and garden equipment.[12] A recent National Research Council report has recommended that the U.S. Environmental Protection Agency (EPA) and others "should undertake the necessary measures to conduct quantitative uncertainty analyses of the mobile source emissions models."[6]

In the area of power plant emissions, Frey and colleagues have developed uncertainty estimates for emissions of hazardous air pollutants and for NO_x emitted by coal-fired power plants.[11,13,14,15] In addition, as part of recent work, methods for quantification of variability and uncertainty have been developed, evaluated, and demonstrated, including the use of Monte Carlo simulation and bootstrap simulation.[16,17,18]

In this paper, quantitative methods for characterizing variability and uncertainty are applied to the source category of stationary natural gas-fueled reciprocating engines. These engines are commonly used, for example, to power natural gas pipeline compressors. In some airsheds, such as for Charlotte, NC, this type of emission source is estimated to be a significant contributor to the total NO_x emission inventory.

OVERVIEW OF METHODS FOR PROBABILISTIC ANALYSIS OF EMISSION FACTORS

Both analytical methods and numerical methods are used for quantification of uncertainty in an unknown quantity.[19] Numerical methods are typically more robust in that they can be applied to a wide range of problems without restrictive assumptions regarding symmetry for probability distributions. In this work, traditional Monte Carlo simulation is employed in uncertainty analysis. The Monte Carlo approach was developed by Stanislaw Ulam and John von Neumann to simulate probabilistic events for military purposes in 1946.[20] Monte Carlo simulation is basically a numerical method for randomly generating sample values from a specified population distribution.

Characterizing Variability in a Data Set

A first step in characterizing variability in a data set is to obtain all relevant data and assess the quality of the data. A judgment must be made that the data are a reasonably representative sample of the population of interest, and that the data are free of significant errors. This step is the same regardless of whether one is developing a point estimate or a probabilistic estimate. This is perhaps the most critical step in the analysis.

A second step is to visualize the data to obtain insight regarding the range, central tendency, and skewness of the data, and any other noteworthy characteristics. A method often employed for this purpose is to plot the data as an empirical cumulative distribution function (CDF). Methods for plotting empirical CDFs are described by Cullen and Frey and by Frey *et al*.[11,19]

One limitation of an empirical CDF is that there is no extrapolation beyond the range of observed data. Thus, for small data set, the real range of variability may be underestimated because variation in samples observed may be much narrower than that in the actual population. Fitting parametric probability distributions has benefits over the use of empirical distribution in that they can provide a plausible means for extrapolating to the unobserved tails of the unknown population distribution.[21] Parametric distributions also tend to have an underlying theoretical basis. In this study, Normal, Lognormal, Gamma, and Weibull distributions are considered. After choosing a candidate parametric distribution that is judged to offer the best fit to the data, the next step is to estimate its parameters based upon the observed data. There are several methods for estimating distribution parameters**.**[19] No method is necessarily the best one to use in all situations. Both Maximum Likelihood Estimation (MLE) and Method of Matching Moments (MoMM) were used in this work.

Characterizing Uncertainty

In this section, a bootstrap simulation method is presented for characterizing uncertainty in any statistic estimated based upon a fitted parametric distribution.

The objective of bootstrap simulation is to numerically simulate sampling distributions for statistics. The main assumption in bootstrap simulation is that the probability distribution estimated from the observed sample of data is the best estimate of the true but unknown population distribution. Given an assumed population distribution, the effects of random sampling from the population distribution are simulated. Specifically, a synthetic data set, known as a *bootstrap sample*, is sampled at random from the assumed population distribution using Monte Carlo simulation. The bootstrap sample has the same number of data points as the original sample. The values of the samples in the bootstrap sample are one possible alternative random realization of the original data set. During bootstrap simulation, a large number of bootstrap samples are simulated, typically 500. For each bootstrap sample, one or more statistics of interest may be calculated, such as the mean. A statistic calculated from a bootstrap sample is referred to as a *bootstrap replication* of the statistic, and there will be random variation in the bootstrap replications. The 500 values of the bootstrap replicates of the statistic can be used to describe a sampling distribution of the statistic. From the sampling distribution, a confidence interval for the statistic can be inferred.

A key advantage of bootstrap simulation for estimation of confidence intervals is that no restrictive assumptions are required regarding normality, as is required to develop confidence intervals using common analytical methods. Thus, bootstrap simulation can be used on a wide variety of problems. The confidence intervals represent lack of knowledge regarding the true values of the statistics being estimated.

NATURAL GAS-FUELED RECIPROCATING ENGINES

Natural gas-fueled reciprocating engines are commonly used to provide mechanical shaft power to drive compressors, such as those used in natural gas pipelines.[22,23] These engines are classified based upon three major designs: (1) 2-cycle lean burn, also referred to as 2-stroke lean burn (2SLB); (2) 4-stroke lean burn (4SLB); and (3) 4-stroke rich burn (4SRB). Natural gas-fueled engines typically emit nitrogen oxides (NO_x), carbon monoxide (CO), and hydrocarbons (HC). Control technologies for natural gas-fueled engines are primarily aimed at reducing NO_x emissions. Emission factors for natural gas-fueled engines have been published by EPA in report number AP-42.[3] Until recently, emission factors for this source category were based upon an October 1996 update to AP-42.[22] However, a more recent update was published in July

2000.[23] The July 2000 version is based upon a different data set than the October 1996 version. The October 1996 data set involves market-share weighted data for NO_x and TOC uncontrolled emission factors, whereas the July 2000 data are assumed to be equally-weighted. To demonstrate a range of analysis methods, both sources of data are included in this study. This study focuses on NO_x and TOC emission factors, because these two pollutants are the most significant precursors to tropospheric ozone formation.

October 1996 Version of Natural Gas-fueled Engine AP-42 Emission Factors

The analysis of the October 1996 version is focused upon lean burn engines, because these engines have high emission rates and are present in an airshed (for Charlotte, NC) that is the subject of a case study in related work. The specific emission sources for which uncertainty in average emission factors were quantified include: (1) 2SLB uncontrolled engines; (2) 2-stroke "clean burn" controlled lean burn engines (2SCB); (3) 2-stroke pre-combustion chamber (PCC) controlled lean burn engines (2SPCC); and (4) 4SLB uncontrolled engines. For other control options, apparently only one data point was used by EPA to estimate emission factors.[24] Therefore, other control options were not analyzed statistically.

For the 2SLB and 4SLB uncontrolled engines, only average emissions data for selected manufacturers were available. In addition, the market share for each manufacturer, in terms of the percentage share of installed capacity, was reported. As an example, the data set for 2SLB engines is given in Table 1. No market share is available for the “Clean Burn” and PCC controlled engines.

The uncontrolled engine emission factors were assigned a data quality rating of “A” by EPA because they judged that the quantity and quality of the original test data were good and generally well documented, and that the engine types and population profile were known. The Clean Burn and Pre-Combustion Chamber controlled engine emission factors were rated as “C,” based on a judgment that the test data were of “A” quality, but that the amount of data was limited.[24]

July 2000 Version of Natural Gas-fueled Engine AP-42 Emission Factors

After the October 1996 version was published, EPA initiated efforts to gather additional emissions data for combustion sources, including stationary reciprocating reciprocating engines. EPA decided to base the emission factors for natural gas-fueled engines on original emissions source test data.[25] The July 2000 emission factors are only for uncontrolled engines. However, the uncontrolled NO_x emission factors have been refined by estimating emissions separately for two different load ranges. EPA has made publicly available the data used to develop the new emission factors in a Microsoft Access database at the EPA TTN web site.[26] A summary of the average emission factor calculated from the data base and of the emission factors reported in AP-42 is given in Table 2. In some cases, it was possible to exactly reproduce the EPA emission factor. However, in other cases it was not possible, and the differences could not be reconciled because of lack of complete documentation by EPA and its contractor regarding how the emission factors were actually calculated.

Two alternative procedures were used to estimate emission factors from the database. In one procedure, referred to in Table 2 as “ungrouped,” each data point in the database was given equal weight, even if some of the data represent repeated measurements of the same engine. In the other procedure, referred to as "grouped," all data for a single engine were averaged, and only the average value for each engine was used to calculate an average emission rate. Of the six emission factors shown in Table 2, it appears that for two of them (2SLB NO_x, both load ranges) it is possible to exactly recalculate the AP-42 emission factor from the available data using the "ungrouped" approach. For both of the TOC emission factors it is possible get a very close approximation to the AP-42 value using the ungrouped approach. For the remaining two emission factors (4SLB NO_x, both load ranges), it is not possible to get a reasonable approximation to the AP-42 value using either approach. After consultation with EPA, it was decided to remove test data collected by Colorado State University from the data set for the 4SLB case, since the CSU test results were more than an order-of-magnitude less than that for the other tests and may have been from a controlled, rather than an uncontrolled, engine. After removing the CSU tests, the ungrouped average is calculated to be 4.40 lb/mmBTU and the grouped average is calculated to be 4.02 lb/mmBTU. The grouped average is very close to the AP-42 value of 4.08 lb/mmBTU for 4SLB engine NO_x emission during 90-105% load operation.

The emission factors of the uncontrolled 2SLB engines were assigned a quality rating “A,” and the emission factors of the uncontrolled 4SLB engines were assigned a quality rating of “B.”[22] However, no explanations regarding the specific basis for these ratings are provided.

QUANTIFICATION OF VARIABILITY AND UNCERTAINTY IN EMISSION FACTORS

Two sets of case studies are presented. In the first case study, each data point is assumed to be an equally likely random sample from the total population of emission sources. This type of case study applies to all of the emission factor data except for the October 1996 version uncontrolled 2SLB and 4SLB engine data, which are weighted by market share and described separately.

Equally-Weighted Randomly Sampled Data

In many cases, emission factor data are available for a sample of engines, representing different manufacturers, engine models, engine ages, and applications. In developing an emission factor, a judgment is made to group data from various specific engine measurements together because of similarities in engine design and operation. For example, expert judgment could be used as a basis for estimating the market share of each particular make and model of engine. In the absence of information, a common default assumption is to assume equal weight among the available data. Of course, this assumption could, and is likely to, be wrong. At the same time, there may not be an empirical basis to justify other assumptions. Key assumptions in an analysis should be evaluated when interpreting the results of the analysis. Therefore, although equal weight for each data point is assumed, later this assumption will be critiqued.

Another factor that must be considered is how to handle replicate data. The available data sets include, in some cases, repeated measurements on the same engine. For example, in the case of the July 2000 data set for uncontrolled NO_x emissions from 4SLB engines operated at 90 percent to 105 percent load, there are 25 data points available from measurements on only 5 engine models. Repeated measurements on the same engine provide an indication of intra-engine variability in emissions. However, in calculating an emission factor, the objective is to quantify inter-engine variability in emissions for purposes of estimating the population distribution for variability within the source category. Therefore, it is necessary to prepare a data set representative of inter-engine variability. The approach taken here is to use an average value for repeated measurements of an individual engine as the representative emission rate for that engine, and to analyze the inter-engine variability in which each engine is represented by either one data point, if only one measurement is available, or the average of the available data, if repeated measurements are available.

The inter-engine variability in emissions for the uncontrolled 4SLB engines is shown graphically in Figure 1. Of the several types of parametric distributions evaluated, the Gamma distribution estimated using MoMM offered the best fit to the four data points. Bootstrap simulation was used to estimate confidence intervals for the CDF of the fitted parametric distribution. With only four data points, the confidence intervals are relatively wide. For example, the 95 percent confidence interval for the median, or 50th percentile of the distribution, is from 2.3 lb/10^6 BTU to 5.7 lb/10^6 BTU, which is nearly as wide as the range of the observed data. The mean emission estimate obtained from the fitted distribution is 4.1 lb/10^6 BTU. The 95 percent confidence interval for the mean is from 2.5 lb/10^6 BTU to 6.1 lb/10^6 BTU, corresponding to a range of minus 39 percent to plus 49 percent.

An important characteristic of the confidence intervals of the mean, or of any other statistic, estimated based upon bootstrap simulation is that they need not be symmetric. With a very small data set of only four data points, and with a positive skewness in the data set, the confidence interval on the mean is expected to be positively skewed. Therefore, the asymmetry of the confidence interval for the mean NO_x emission factor from 4SLB engines is expected. Because of the small number of data points and the wide range of variability of the data, the confidence interval is expected to be relatively wide, as it is in this case.

The adequacy of the fitted distribution can be evaluated, at least in part, by identifying what proportion of the data are contained with the confidence intervals of the CDF. On average, if the fit is a good one, half of the data should be enclosed within the 50 percent confidence interval, 90 percent of the data should be enclosed within the 90 percent confidence interval, and 95 percent of the data should be enclosed within the 95 percent confidence interval. In Figure 1, three of the four data points are contained within the 50 percent confidence interval, and all of the data are enclosed by the 90 percent confidence interval. This suggests, though cannot prove, that the Gamma distribution is an acceptable fit to the data.

Unequally-Weighted Data

In this section, an example case study is presented based upon emissions data that are not equally weighted. These data are from Table 1 for uncontrolled 2SLB engines, based upon the October 1996 version of AP-42. The five emissions values are shown in Figure 2 as an empirical CDF, along with three parametric distributions that have been fit to the data.

Because each of the five emissions values has a different market share-based weight, the method for fitting distributions to the data had to be modified compared to when data have equal weight. The approach taken here was to use 100 synthetic data points as a basis. The use of 100 basis data points allows for emission values to occur repeatedly in proportion to their market share. A portion of these 100 data points were assigned the emission factor associated with an engine, in proportion to the market share of that engine. For example, the Clark engines have 36 percent of the market share; therefore, 36 of the 100 basis data points were assigned the Clark engine emission value of 2.64 lb/10^6 BTU. Parametric distributions were fit to the 100 basis data points.

The comparison of the fitted distributions in Figure 2 suggests that the Weibull distribution may provide the best fit to the data. The Weibull distribution provides the best fit in the central portion of the distribution, and appears not to have as "heavy" of a tail at the upper end of the distribution. For comparison purposes, both the Weibull and Lognormal distributions are included in the bootstrap simulation analyses.

During bootstrap simulation, each simulated data point has equal weight. However, because the parametric distributions were fit to market share-weighted data, the shape of the parametric distributions reflects the frequency with which data should be sampled in different emission ranges. For example, the steepness of the fitted CDF in the range from approximately 2 lb/10^6 BTU to 3 lb/10^6 BTU means that there is a high probability that random samples of emissions will occur in this range, corresponding to the three engines that have the largest combined market share. In contrast, there is comparatively little probability that emissions values will be sampled for the two engines that, together, comprise only five percent of the total market share.

The results of the bootstrap simulation with the Lognormal distribution are given in Figure 3. It appears that the 95 percent confidence interval encloses the empirical distribution of the data. However, the confidence intervals are very wide, and there appear to be biases in the fit. For example, the central range of the empirical distribution coincides with the high side of the confidence intervals, while the lower and upper tails of the empirical distribution coincide with the low side of the confidence interval. The apparent biases in the fit, and the wideness of the intervals, suggest that the Lognormal is not a particularly good distribution to use in this case.

The results of the bootstrap simulation with the Weibull distribution are given in Figure 4. These results imply more consistency between the assumed parametric distribution and the empirical distribution of the original data. In particular, the empirical distribution appears to be reasonably well enclosed by the 90 percent confidence interval, and the width of the confidence interval is much narrower compared to the Lognormal case, without compromising the apparent goodness-of-fit. Therefore, the Weibull distribution is selected over the Lognormal distribution as a more appropriate basis for estimating uncertainty in the mean. The choice of parametric distribution influences the estimated confidence interval for the mean. The 95 percent confidence interval for the mean is 2.14 to 3.38 lb/10^6 BTU based upon the Lognormal distribution, 2.25 to 3.26 lb/10^6 BTU based upon the Gamma distribution, and 2.39 to 2.99 lb/10^6

BTU based upon the Weibull distribution. Of these three, the Weibull distribution leads to the narrowest estimate of the confidence interval.

In order to compare the influence of the market-share factor on the quantification result. This data set is also analyzed as equally weighted data. The result of bootstrap simulation with fitted Weibull distribution is given in Figure 5. Comparing Figure 5 with Figure 4, if unequally-wieghted data are treated equally, the mean of the bootstrap means is 1.95 lb/10^6 BTU, which is 28 percent smaller than the value if treated unequally; the 95 percent confidence interval is 1.22 to 2.75 lb/10^6 BTU, which 160 percent wider than if the data are treated unequally; the lower bound and the upper bound of 95 percent confidence interval is 49 percent smaller and 8 percent smaller, respectively, than the values if treated unequally. The two engines with low emission comprise only 5 percent of total market share. If they are improperly given the same weight as the other three engines, there will be an underestimation of the overall mean emission. Therefore, properly reporting the market share along with the emission factor data is important for accurately estimating the emission factors.

Summary of Probabilistic Estimation Results for AP-42 October 1996 Version and July 2000 Version Emission Factors

Table 3 gives quantified uncertainties in NO_x emission factors for natural gas engines. Table 4 gives quantified uncertainties in TOC emission factors for natural gas engines. The probabilistic estimates presented in Table 3 and 4 are based on October 1996 AP-42 data. The quantified uncertainties in NO_x emission factors for different operation loads based on July 2000 AP-42 data are presented in Table 5. Table 6 gives quantified uncertainties in TOC emission factors based on July 2000 AP-42 data. The summary tables indicate that the 95 percent range of uncertainty in the mean emission factor ranges from as low as approximately plus or minus 10 percent to as high as minus 80 to plus 180 percent. The range of uncertainty is influenced by a combination of the sample size and the range of variability in the data. Smaller sample sizes and/or larger inter-engine variability in the data will tend to contribute to wider ranges of uncertainty in the estimated mean emission factor.

CONCLUSIONS

This paper demonstrates the successful application of quantitative probabilistic analysis to emission factor case studies, based upon the example of stationary natural gas-fueled reciprocating engines. The method employed is based upon characterization of uncertainty based upon random sampling error. The method includes: (1) development of a database; (2) visualization of the data using empirical CDFs; (3) evaluation of alternative parametric probability distributions fitted to the data; (4) bootstrap simulation to characterize confidence intervals in the fitted CDF; (5) selection of a judged best fit distribution based upon bootstrap simulation results; and (6) quantification of uncertainty in the mean based upon the bootstrap sampling distribution for the mean.

The probabilistic method was applied to several different types of analyses, including: (1) quantification of inter-engine variability in emissions and uncertainty in the mean for unequally weighted data points; and (2) quantification of inter-engine variability in emissions and uncertainty in the mean for equally weighted data points. The range of inter-engine variability in emissions suggests that the weights assigned to each engine emission estimate can significantly affect the estimate of the mean emission rate. Thus, the assumption of equal weighting of emissions data, as is often made, is likely to be a strong assumption in many cases and, therefore, can be a significant factor biasing emission factor estimates.

The estimates of uncertainty in the mean are often asymmetric, indicating that skewness regarding observed variability in inter-engine emissions can lead to skewness in the estimate of uncertainty in the mean. Conventional analytical methods based upon normality assumptions can lead to errors in the uncertainty estimate. The mean values estimated from the probabilistic analysis differ in some cases from the mean values estimated directly from the data because parametric probability distributions allow for interpolation within the range of observed data and for extrapolations beyond the range of observed data. For small data sets, it is unlikely that the observed sample of data truly includes the minimum and maximum possible values. On this basis, extrapolation is warranted.

Although three parametric distributions were typically evaluated, most often the Weibull distribution was found to provide a good fit to the data. The Weibull may take on many shapes, including negatively skewed, symmetric, or positively skewed. Furthermore, the Weibull distribution tends to be less "tail-heavy" than the other two, and often provides a better empirical fit to the data for these reasons.

The quantitative analysis demonstrated here focuses on one important source of uncertainty. The range of uncertainty associated with random sampling error was found to be as large as minus 80 percent to plus 180 percent, and in most examples was greater than plus or minus 20 percent. Some other sources of uncertainty, such as potential lack of representativeness of the test cycles used in the measurements, or potential lack of representativeness of the sample of engines, are difficult to evaluate quantitatively. Therefore, it is recommended that qualitative methods for identifying sources of uncertainty *also* be used. However, there is not a direct relationship between the qualitative data rating and the range of uncertainty in the emission factor. Therefore, we do not recommend that data quality ratings be used to make inferences regarding quantitative ranges of uncertainty.

A significant difficulty encountered in this study was the lack of documentation of the calculation methods for the July 2000 AP-42 emission factors. Complete documentation should include enough information so that others can reproduce the calculations and results. Therefore, we recommend that EPA report the complete calculation method used for each emission factor. With the growing recognition of the importance of quantitative uncertainty analysis, it will be important for EPA and others to routinely report data regarding variability and uncertainty in emission factors.

ACKNOWLEDGEMENTS

This work was supported by U.S. Environmental Protection Agency STAR Grants Nos. R826766 and R826790. This paper has not been subjected to any EPA review. Therefore it does not necessarily reflect the views of the agency, and no official endorsement should be inferred.

We thank Mr. Roy Hunter of U.S. EPA in Research Triangle Park, NC, for his assistance and recommendations regarding data sources.

REFERENCES

1. Bogen, K. T.; Spear, R. C. "Integrating Uncertainty and Interindividual Variability in Environmental Risk Assessment", *Risk Anal.* 1987, 7, 427-436.
2. Hattis, D.; Burmaster, D. E. "Assessment of Variability and Uncertainty Distributions for Practical Risk Analyses", *Risk Anal.* 1994, 14, 713-730.
3. EPA. *Compilation of Air Pollutant Emission Factors, AP-42, Fifth Edition*; US Environmental Protection Agency, Research Triangle Park, NC, 1995.
4. Beck, L.L.; Wilson, D. "EPA's Data Attribute Rating System". In *Proceedings, Emission Inventory: Planning for the Future*, Air & Waste Management Assoc.: Pittsburgh, PA, 1997.
5. EPA. *Guiding Principles for Monte Carlo Analysis*, US Environmental Protection Agency, Research Triangle Park, NC, 1997; EPA/630R-97/001
6. National Research Council. *Modeling Mobile-Source Emissions*; National Academy Press; Washington, DC, 2000.
7. National Academy of Science. *Science and Judgement in Risk Assessment*; National Academy Press; Washington, DC, 1994.
8. IPCC National Greenhouse Gas Inventories Program, *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)
9. Kini, M.D.; Frey, H.C. "Probabilistic Evaluation of Mobile Source Air Pollution"; Prepared for Center for Transportation and the Environment by North Carolina State University, Raleigh, NC, 1997.
10. 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, Prepared for Coordinating Research Council by ENVIRON International Corp, Novato, CA, 1999.
11. Frey, H.C.; Bharvirkar, R.; Zheng, J. "Quantitative Analysis of Variability and Uncertainty in Emissions Estimation"; Prepared for the U.S. Environmental Protection Agency by North Carolina State University, Raleigh, NC, 1999. (available at http://www4.ncsu.edu/~frey/Frey_Bharvirkar_Zheng_99.pdf).
12. Bammi, S.; Frey, H.C. "Quantification of Variability and Uncertainty in Lawn and Garden Equipment NO_x and Total Hydrocarbon Emission Factors", In *Proceedings, 94th Annual Meeting of the A&WMA*, Air & Waste Management Assoc.: Pittsburgh, PA, 2001.
13. Frey, H. C.; Rhodes, D. S. "Characterizing, Simulating, and Analyzing Variability and Uncertainty: An Illustration of Methods Using an Air Toxics Emissions Example", *Human Health and Ecological Risk Assessment*, 1996, 2(4), 762-797.
14. Frey, H.C.; Bharvirkar, R. "Quantification of Variability and Uncertainty: A Case Study of Power Plant Hazardous Air Pollutant Emissions", *The Risk Assessment of Environmental and Human Health Hazards: A Textbook of Case Studies*; D. Paustenbach, Ed., John Wiley and Sons; New York. 2001.

15. Rhodes, D.S.; Frey, H.C. “Quantification of Variability and Uncertainty in AP-42 Emission Factors Using Bootstrap Simulation”, In *Proceeding, Emission Inventory: Planning for the Future*, Air and Waste Management Assoc.: Pittsburgh, PA, 1997; pp 147-161.
16. Frey, H.C.; Rhodes, D.S. “Characterization and Simulation of Uncertain Frequency Distributions: Effects of Distribution Choice, Variability, Uncertainty, and Parameter Dependence”, *Human and Ecological Risk Assessment*, 1998, 4(2), 423-468.
17. Frey, H. C.; Burmaster, D. E. “Methods for Characterizing Variability and Uncertainty: Comparison of Bootstrap Simulation and Likelihood-Based Approaches”, *Risk Analysis*, 1999, 19(1), 109-129.
18. Zheng, J.; Frey, H.C. “Quantitative Analysis of Variability and Uncertainty: An Illustration of Methods Using Mixture Distributions”, In *Proceedings, 94th Annual Meeting of the A&WMA*, Air & Waste Management Association, Pittsburgh, PA, 2001.
19. Cullen, A. C.; Frey, H. C. *Probabilistic Techniques in Exposure Assessment: A Handbook for Dealing with Variability and Uncertainty in Risk Analysis*; Plenum Press; New York, 1999.
20. Rugen, P.; Callahan, B. “An Overview of Monte Carlo A fifty Year Perspective”, *Human Health and Ecological Risk Assessment*, 1996, 2(4), 671-680.
21. Eastern Research Group. “Report of the Workshop on Selecting Input Distributions for Probabilistic Assessments”; EPA/630/R-98/004, Prepared for Risk Assessment Forum by Eastern Research Group, Morrisville, NC, 1999.
22. EPA. “Volume I: Stationary Point and Area Sources, Chapter 3.2 Heavy-duty Natural Gas-fueled Pipeline Compressor Engines And Turbines”, *Compilation of Air Pollutant Emission Factors, AP-42, Fifth Edition*; U.S. Environmental Protection Agency; Research Triangle Park, NC, 1996.
23. EPA. “Volume I: Stationary Point and Area Sources, Chapter 3.2 Natural Gas-fueled Reciprocating Engines”, *Compilation of Air Pollutant Emission Factors, AP-42, Fifth Edition*; U.S. Environmental Protection Agency; Research Triangle Park, NC, 2000.
24. Eastern Research Group. “Report on Revisions to 5th Edition AP-42, Section 3.2 Heavy-duty Natural Gas-fueled Pipeline Compressor Engines”; Prepared for U.S. Environmental Protection Agency by Eastern Research Group, Morrisville, NC, 1996.
25. Alpha-Gamma Technologies, Inc. “Emission Factor Documentation for AP-42 Section 3.2, Natural Gas-fueled Reciprocating Engines”, Prepared for U.S. Environmental Protection Agency by Alpha-Gamma Technologies, Inc., Raleigh, NC, 2000.
26. EPA emissions data web site, located December 2000. (http://www.epa.gov/ttn/chief/ap42/ch03/related/c03s02.html).

KEY WORD

Emission factors, NO_x, TOC, Uncertainty, Variability, Bootstrap simulation, Natural Gas Engine.

Figure 1. Comparison of Empirical Cumulative Distribution of Average Uncontrolled 4-SLB Engine, 90-105% load, NO_x Emissions, fitted Weibull distribution, and Bootstrap Simulation Confidence Intervals, Based Upon July 2000 AP-42 Data.

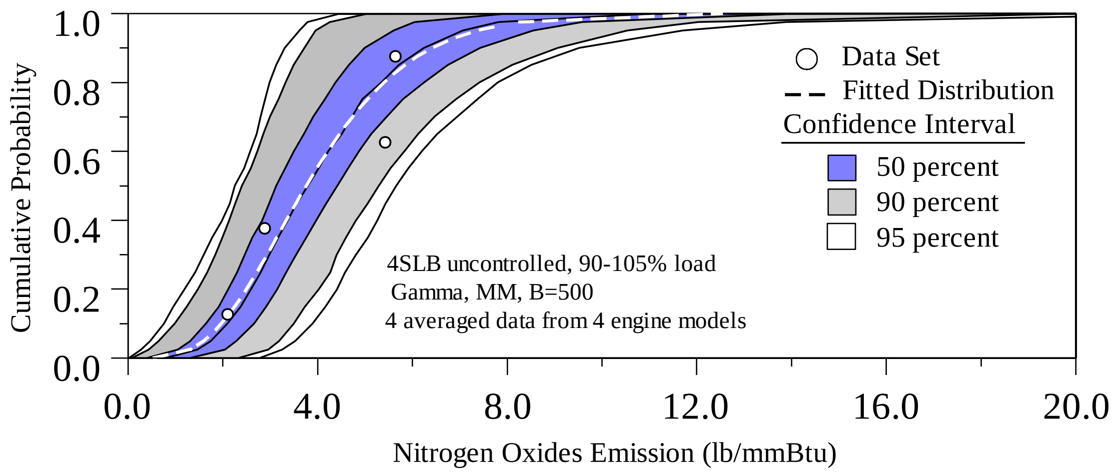

Figure 2. Empirical Distribution and Fitted Parametric Distributions for Market-Share Weighted NO_x Emissions Rates for Uncontrolled 2-Cycle Lean Burn Engines Based Upon October 1996 AP-42 Data

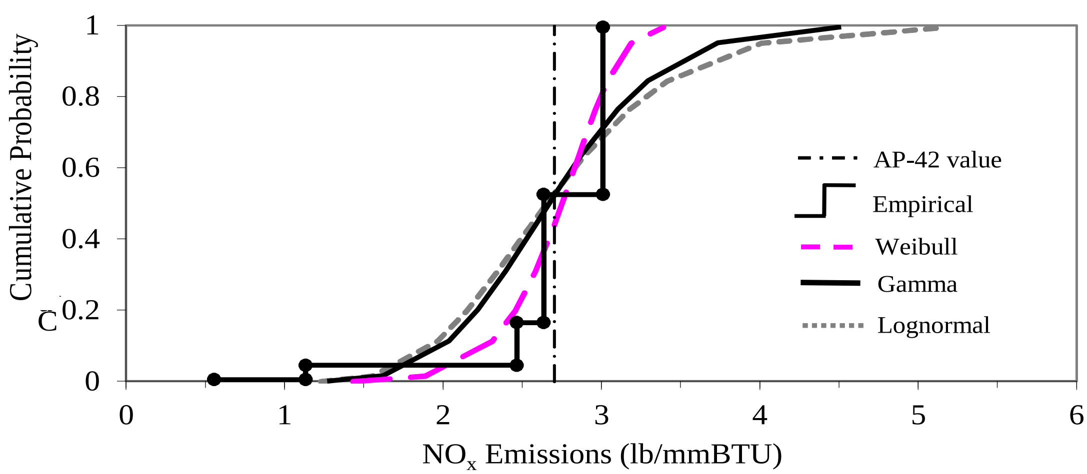

Figure 3. Comparison of the Empirical Distribution Bootstrap Simulation Results Based Upon a Lognormal Distribution for Market-Share Weighted NO_x Emissions Rates for Uncontrolled 2-Cycle Lean Burn Engines Based Upon October 1996 AP-42 Data

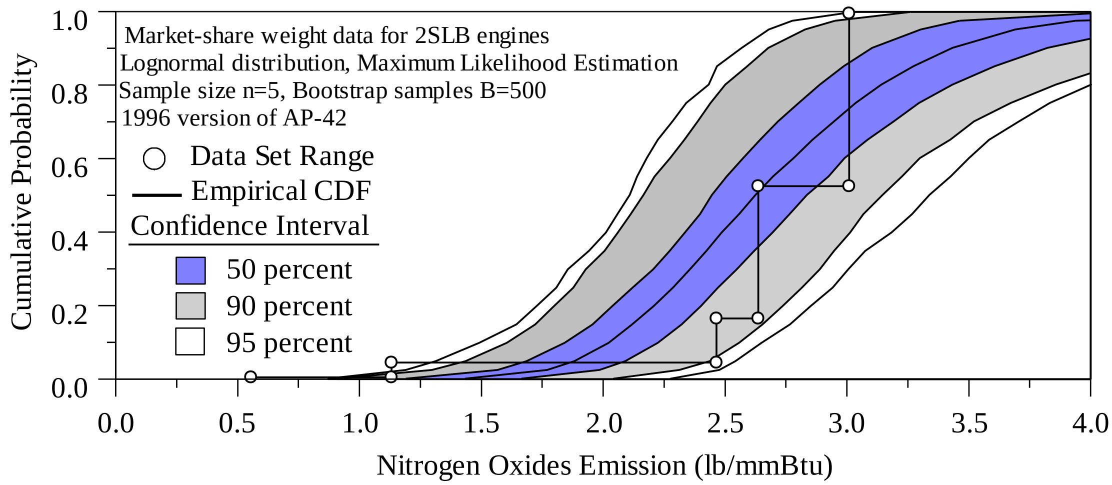

Figure 4. Comparison of the Empirical Distribution Bootstrap Simulation Results Based Upon a Weibull Distribution for Market-Share Weighted NO_x Emissions Rates for Uncontrolled 2-Cycle Lean Burn Engines Based Upon October 1996 AP-42 Data

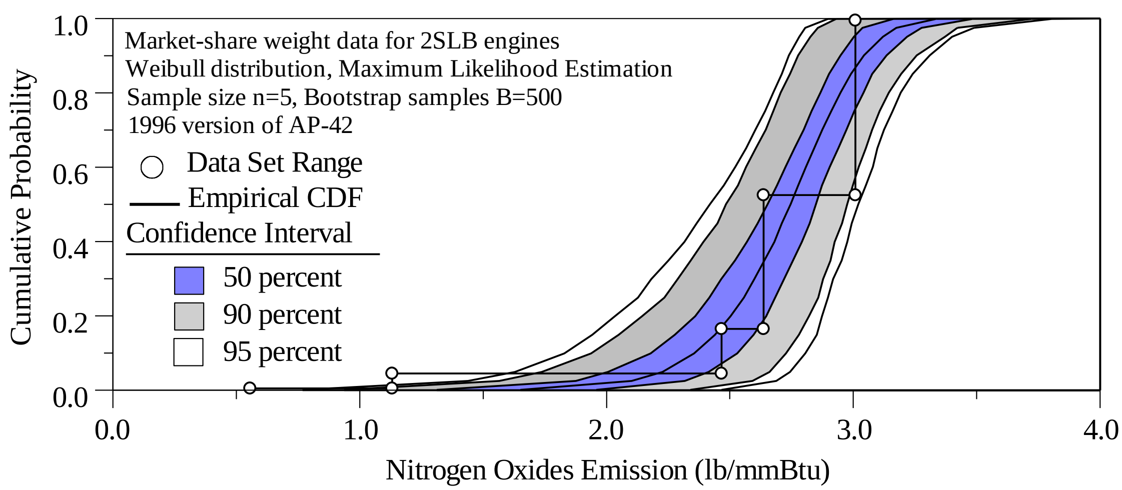

Figure 5. Comparison of the Empirical Distribution Bootstrap Simulation Results Based Upon a Weibull Distribution, Market-Share Weighted NO_x Emissions Rates, Treated as Unweighted data, Uncontrolled 2-Cycle Lean Burn Engines Based Upon October 1996 AP-42

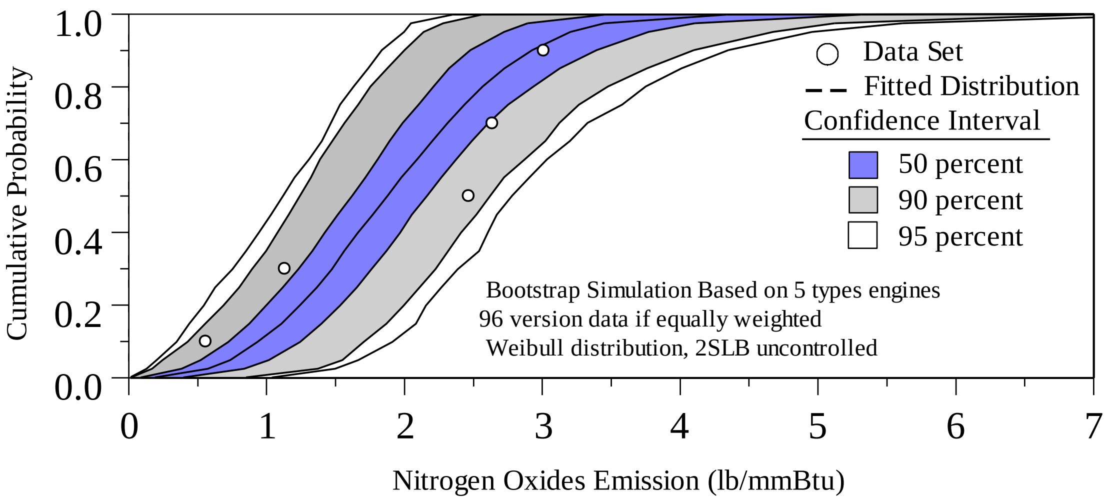

Table 1. Emissions data for Uncontrolled Natural-Gas Fueled 2-Stroke Lean Burn Engines.[21]

MAKE	NO_x Emissions ($lb/10^6$ BTU)	TOC Emissions ($lb/10^6$ BTU)	Ratio of total installed capacity (%)
Ajax	1.132	4.318	4
Clark	2.636	1.703	36
CB	3.009	1.164	47
Fairbanks-Morse	0.556	1.220	1
Worthington	2.466	1.618	12
Weighted average	2.710	1.539	

Table 2. Comparison Between EPA NO_x Emissions Database and Documentation of AP-42 Emission Factors for Uncontrolled 2SLB and 4SLB Engines Based Upon July 2000 Version of AP-42.[22]

Engine Type	Pollutant	Engine Load	Average Calculated from Database[a] ($lb/10^6$ Btu)	AP-42 Emission Factor ($lb/10^6$ Btu)	Comments on Documentation[b]
2SLB	NO_x	90 to 105%	3.17 (ungrouped), 3.05 (grouped)	3.17	34 test data are used to develop AP-42 emission factor
		< 90%	1.94 (ungrouped), 2.15 (grouped)	1.94	57 test data are used to develop AP-42 emission factor
	TOC[c]	Any load	1.61(ungrouped), 1.49 (grouped)	1.64	24 test data are used to develop AP-42 emission factor
4SLB	NO_x	90 to 105%	2.22 (ungrouped) 3.26 (grouped)	4.08	25 test data are used to develop AP-42 emission factor
		< 90%	0.739 (ungrouped) 1.44 (grouped)	0.847	13 test data are used to develop AP-42 emission factor
	TOC[c]	Any load	1.42(ungrouped), 1.13 (grouped)	1.47	37 test data are used to develop AP-42 emission factor

[a] Two average values were calculated from the available data in the database from the EPA TTN Web Site. The "Ungrouped" averages involve taking the average of all emissions tests for all engines. The "Grouped" averages involve first calculating the average emissions for engines that were tested more than once, and then calculating the average among all engines. For example, if we have 25 test data from 10 engines, the ungrouped average is based upon 25 equally weighted values. In contrast, the grouped average would be based on the 10 average values for each different engine.

[b] The test identification numbers used in the on-line database are documented in Reference 22.

[c] Emission factors are reported on a TOC basis in AP-42. While, they are reported as Total Hydrocarbons (THC) in database.[20,23]

Table 3. 95 Percent Confidence Interval for Mean NO_x Emissions for Natural Gas-fired Reciprocating Lean Burn Engines, Based on October 1996 AP-42 Data.

Engine and Emissions Control Technology	No. of Data	Mean of Data[a]	AP-42 Emission Factor[a]	Fitted Distrib.[d]	Mean of Bootstrap Sample Means[a]	Relative 95% CI on Mean[b]
2SLB, Uncontrolled	5	2.710	2.710	Weibull	2.714	-11.8% to +9.36%
2SLB, Clean Burn	11	0.834	0.834	Lognormal	0.835	-14.1% to +15.4%
2SLB, PCC[c]	20	0.850	0.850	Lognormal	0.840	-23.7% to +28.5%
4SLB, Uncontrolled	4	3.225	3.225	Weibull	3.170	-27.2% to +30.8%

[a]Units are lb/10^6 BTU. [b]Calculated based upon bootstrap simulation results. [c]PCC=Pre-Combustion Chamber
[d]MLE is used for parameter estimation

Table 4. 95 Percent Confidence Intervals for Mean TOC Emissions for Natural Gas-fired Reciprocating Lean Burn Engines, Based on October 1996 AP-42 Data.

Engine and Emissions Control Technology	No. of Data	Mean of Data[a]	AP-42 Emission Factor[a]	Fitted Distrib. [d]	Mean of Bootstrap Sample Means[a]	Relative 95% CI on Mean[b]
2SLB, Uncontrolled	5	1.539	1.539	Weibull	1.549	-36.0% to +42.7%
2SLB, Clean Burn	11	0.767	0.767	Weibull	0.770	-56.1% to +67.5%
2SLB, PCC[c]	20	1.756	1.756	Weibull	1.750	-17.1% to +18.3%
4SLB, Uncontrolled	4	1.261	1.261	Weibull	1.278	-47.6% to +55.7%

[a]Units are lb/10^6 BTU. [b]Calculated based upon bootstrap simulation results. [c]PCC=Pre-Combustion Chamber
[d]MLE is used for parameter estimation

Table 5. 95 Percent Confidence Intervals for Mean Uncontrolled NO_x Emissions for Natural Gas-fired Reciprocating Lean Burn Engines, Based on July 2000 AP-42 Data.

Engine and Load Range	AP-42 Emission Factor[a]	No. of Data	No. of Engines	Fitted Distrib. [d]	Mean of Bootstrap Sample Means[a]	Relative 95% CI on Mean[b]
2SLB, 90% to 105%	3.17	34	11	Weibull	3.05	-24% to +24%
2SLB, < 90%	1.94	24	11	Weibull	2.18	-41% to +46%
4SLB, 90% to 105%	4.08	12	4	Gamma	4.06	-39% to +49%
4SLB, < 90%	0.847	13	5	Weibull	1.81	-80% to +180%

[a]Units are lb/10^6 BTU. [b]Calculated based upon bootstrap simulation results.
[d]MLE is used for 2SLB engine, MoMM is used for 4SLB engine

Table 6. 95 Percent Confidence Intervals for Mean TOC Emissions for Natural Gas-fired Reciprocating Lean Burn Engines, Based on July 2000 AP-42 Data.

Engine and Load Range	AP-42 Emission Factor[a]	No. of Data	No. of Engines	Fitted Distrib. [d]	Mean of Bootstrap Sample Means[a]	Relative 95% CI on Mean[b]
2SLB, Uncontrolled	1.64	57	14	Weibull	1.45	-16% to +18%
4SLB, Uncontrolled	1.47	37	4	Gamma	1.12	-45% to +57%

[a]Units are lb/10^6 BTU. [b]Calculated based upon bootstrap simulation results.
[d]MLE is used for 2SLB engine, MoMM is used for 4SLB engine