In *Proceedings, Annual Meeting of the Air & Waste Management Association*, June 24-28, 2001, held in Orlando, Florida, and published byA&WMA, Pittsburgh, PA

Quantification Of Variability and Uncertainty In Lawn And Garden Equipment NO_x and Total Hydrocarbon Emission Factors

2001-ABSTRACT #580

Sachin Bammi
Research Assistant, Department of Civil Engineering, North Carolina State University

H. Christopher Frey
Associate Professor, Department of Civil Engineering, North Carolina State University

ABSTRACT

Emission factors and emissions estimates are subject both to variability and uncertainty. Variability refers to real differences in emissions among multiple emission sources at any given time or over time for any individual emission source. Variability in emissions can be attributed to variation in: fuel or feedstock composition; ambient temperature; process design; process maintenance; and/or process operation. Uncertainty refers to lack of knowledge regarding the true value of emissions. Sources of uncertainty include: small sample sizes; lack of precision and/or accuracy in measurements; non-representativeness; or lack of data. In this paper, we demonstrate quantitative methods for characterizing both variability and uncertainty and apply the methods to case studies of emission factors for lawn and garden engines. Lawn and Garden engines, which are primarily 2-stroke and 4-stroke gasoline engines of less than 25 hp, are used in a variety of applications including walk behind mowers, lawn tractors, shredders, grinders, blowers, vacuums, string trimmers, chainsaws, hedge trimmers, and others. Data were obtained from emissions testing conducted by others and were analyzed to quantify both variability and uncertainty. The variability in emissions was quantified using empirical and parametric distributions. Bootstrap simulation was used to characterize confidence intervals for the fitted distributions. The 95 percent confidence intervals for the mean emission factors associated with 2S total hydrocarbon (THC) and Nitrogen Oxides (NO_x) emissions in g/hp-hr units were -32% to +38% and -46% to +65%, respectively. The confidence intervals for mean emissions in g/hp-hr units for 4S engines were -38% to +45% for THC and -25% to +38% for NO_x. These quantitative measures of uncertainty convey information regarding the quality of the emission factors and serve as a basis for calculation of uncertainty in emission inventories. The method, example case studies, and benefits of the approach will be presented.

1.0 INTRODUCTION

Emission Inventories (EIs) are an important part of environmental decision making. They are used for temporal emissions trend characterization, emissions budgeting for regulatory and compliance purposes and in air quality models for estimating ambient air quality. Non-quantified errors and biases in the EIs can lead to erroneous conclusions regarding emissions trends, apportionment of emissions, compliance, and the relationship between emissions and the ambient pollutant concentration. Hence, characterization and evaluation of quality of data should be an integral part of emissions inventory work. The National Research Council, in a recent report on modeling of mobile source emissions, has recommended that efforts be conducted to

quantify uncertainties in mobile source emissions.[1] In this paper, a quantitative approach to the characterization of variability and uncertainty is presented as an important aspect in conveying the quality of estimates to analysts and decision makers. This paper focuses on the application of quantitative analysis of variability and uncertainty to a case study for non-road mobile sources.

The 1990 Clean Air Act Amendments specifically directed the U.S. Environmental Protection Agency (EPA) to study the contribution of nonroad engines to urban air pollution, and regulate them if warranted. "Nonroad" is a term that covers a diverse collection of engines, equipment, and vehicles. Also referred to as "off-road" or "off-highway," the nonroad category includes outdoor power equipment, recreational equipment, farm equipment, construction equipment, lawn and garden equipment, and marine vessels.[2]

In the next section, there is a discussion of probabilistic analysis and applications to emission inventories. In the third section, a review of current emission factor estimation methods for Lawn and Garden (L&G) engines is provided. In section four, L&G emission factor data are reviewed and evaluated. In section five, analyses are presented which include comparison of the various test cycles from which data are used to development emission factors, evaluation of possible groupings of the emissions data, and probabilistic analysis of the emission factors.

2.0 PROBABILISTIC ANALYSIS OF EMISSION INVENTORIES

The quantitative approach to probabilistic analysis of both variability and uncertainty employed in this work differs from the conventional approach to emission factor and inventory data quality ratings. Qualitative “A” through “E” ratings are defined and reported by EPA.[3] The Data Attribute Rating System (DARS) is a method for combining data quality scores for both emission factor and activity data to develop an overall quality score for an emission inventory.[4] While DARS can be used to compare quality ratings for EIs, it can neither be used to quantify the precision of an inventory nor to evaluate the robustness of a decision to uncertainty. Other efforts have focused on characterizing the mean and variance of emission estimates and using simplified approaches for combining uncertainties in activity and emission factor data to arrive at an aggregate uncertainty estimate.[5, 6,7] The applications of these approaches suffer from many shortcomings including: restrictive assumptions about the shape of probability distribution models; failure to distinguish between variability and uncertainty estimates; and improperly analyzed small sample data. The probabilistic approach presented here will remedy these shortcomings.

EPA, the National Academy of Sciences (NAS), and others recognize the important role of probabilistic analysis.[8,1] In March 1997, EPA issued “Guiding Principles for Monte Carlo Analysis”.[9] The policy supports “good scientific practices” in quantifying variability and uncertainty. Therefore it is critically important that research be undertaken in the development and application of probabilistic methods. An overview of probabilistic analysis methods is given by Cullen and Frey.[10] The recent NRC recommendations regarding uncertainty analysis for mobile sources also motivate this work.[1] There have been some efforts to quantify uncertainty in mobile sources, focused on highway vehicles.[11,12,13] In addition, there has been work to quantify uncertainties in emissions of stationary point and area sources.[14,15,13,16] In this paper, a methodology for simultaneous characterization of both variability and uncertainty is presented, with applications to non-road emission sources.

3.0 A REVIEW AND EVALUATION OF CURRENT EMISSION FACTOR DATA FOR L&G EQUIPMENT

The original objective of the data collection activity was to collect all of the test data that was used by the EPA in coming up with the emission factors for the L&G nonroad emission source category and then reproduce those emission factors. However, the data to do this was not easily available. EPA uses emission factors in the NONROAD emission inventory model. Information regarding the specific emission factor data used as input to the model are not readily available. In particular, the emissions data are from a variety of reports, a number of which are not available in the open literature and could not be obtained after several months of searching. Given the unavailability of a complete data set and the incomplete documentation of the current version of the NONROAD model, it was decided instead to focus on obtaining a reasonable data set and analyzing these data to develop relative measures of uncertainty. Relative measures of uncertainty, such as plus or minus percentage ranges on the estimated mean value; provide insight into the level of uncertainty anticipated in any emission factor based upon a similar dataset used as input to the NONROAD model.

All of the emission factor data that were collected and studied for the purposes of this project came from technical reports of the California Air Resources Board (CARB)[17], and contract reports prepared by Southwest Research Institute (SwRI)[18,19,20,21,22], and papers published in the Journal of the Air and Waste Management Association (AWMA)[23], and by the Society of Automotive Engineers (SAE).[24,25,26,27,28] Units in which emission factor data were reported were either: (1) grams per hour (g/hr); (2) grams per brake horsepower-hour (g/hp-hr); and/or (3) grams per kilowatt-hour (g/kw-hr). For evaluation purposes, some of these data were converted to gram per gallon (g/gallon) units, assuming that the specific gravity of gasoline is 0.75.[29]

The available L&G emission factor data were mostly generated by using the SAE J1088 or CARB J1088 test procedures.[18,21] These test methods involve measuring emissions during a standard series of "mode" tests, in which each mode represents a typical type of steady-state operation (e.g., idle, full load). Each method has a specific number of modes and a specific method for weighting the modes to estimate an overall average emission rate. These test procedures are further described later in this paper. The SAE and CARB variations of the J1088 procedure are the most popular of the various test methods that have been citied in the literature. Some investigators also used some other test methods. However, many of the other test methods use some modes that are similar to those in the J1088 tests. There was no data found regarding quality ratings for the emission factors reported based upon the various tests identified in the literature review. There were also no indicators of variability and uncertainty in the reported emission factors.

In addition to references that contained emissions testing data, two other reports form the EPA were reviewed, including the Nonroad Engine and Vehicle Emissions Study Report and documentation of EPA's most recent NONROAD model.[30,31] These two reports provide general information regarding the types of data and categories used by EPA to calculate emission factors. However, the emission factor data needed for purposes of statistical analysis were not available in these two reports. Instead, emissions measurements databases had to be developed based upon the emission test reports described above.

In addition to identifying sources of data, the literature review also involved identifying what categories of L&G emission sources EPA has assumed in estimating emission factors. The next three subsections discuss alternative groups of L&G emission factor data.

3.1 Application vs. Power Rating

EPA, in the NEVES study, grouped emission factors for L&G engines by application. EPA did not report any difference between emission factors for different sizes within the same application. On the other hand, there were other nonroad emissions models, such as CARB's OFFROAD Model, which had emission factors for L&G engines grouped by engine size. In OFFROAD the emission factors were grouped by engine size using their horsepower rating. The latest NONROAD model that EPA is developing groups emission factors for L&G engines primarily by power level, but it also uses "application specific" emission factors where there is "sufficient" information to justify their use.[31]

3.2 Two-Stroke vs. Four-Stroke

EPA and CARB have categorized emission factors with respect to two-stroke versus four-stroke classifications.[17,30] Two-stroke engines are used in equipment such as chainsaws, string trimmers, blowers, vacuums, small generators and pumps, and in walk behind mowers. Two-stroke engines dominate applications requiring a lightweight, high power engine. Four stoke engines dominate the "general utility" equipment category and are also used in the majority of L&G equipment including walk behind mowers, riding mowers and tractors, tillers, grinders, and other equipment where the implement is supported by wheels rather than hand-held. Hydrocarbon emissions from two-stroke engines range from 10 to 15 times higher than for similarly powered four-strokes. Four stoke engines contribute approximately eight times the NO_x emissions as two-stroke engines.[17]

3.3 Overhead Valve (OHV) Design vs. Side Valve (LHV) Design

Four stroke engines include two major design variations: (1) over head valve (OHV); and (2) side valve (LHV). OHV and LHV are devices that open and close the combustion chamber of an internal-combustion engine to admit the fuel-air mixture or exhaust the gases. The intake and the exhaust ports, in the LHV design, are positioned on one side of the cylinder. In case of the OHV design both valves are located directly over the piston. CARB has categorized 4 stroke engine emission factors as either OHV or LHV. Compared to similarly powered LHV engines, OHV engines have lower emissions, better fuel economy, and longer life.[17]

3.4 Handheld vs. Non-Handheld

Handheld engines are those that need to be held in hand when being used, such as string trimmers or chainsaws. Non-Handheld engines are those that are not held in hand while in use, such as walk behind mowers and garden tractors. Most of the handheld engines are two-stroke engines, with a few exceptions. Similarly most of the non-handheld engines are four-stroke engines. Hence the difference in the emissions between these two categories is similar to the difference in emissions between 2-stroke and 4-stroke engine categories. The CARB J1088 test procedure proposes different test cycles for handheld versus non-handheld applications for a given type of engine (e.g., 2-Stroke).

4.0 DATA ANALYSIS

Data collected from the literature review were analyzed. First, the reproduction of the reported emission factors is discussed followed by comparisons between the various test procedures that were used to develop the reported emission factors. Finally the results of statistical analysis of the data to ascertain statistically significant groupings among the emission factors are presented.

4.1 Reproducing The Reported Emission Factors

To better understand the emission factor development process, the emission factors for THC and NO_x that were reported in the SwRI reports were recalculated from the modal engine test data given in these reports. This was possible with the SwRI reports because they contained the actual test mode data and not just the weighted averages of all modes. It was possible to recalculate the reported average emission factors for each test to within three or four significant digits. The units with emission factors were reported varied among the data sources. For example, g/hp-hr units are reported by CARB, whereas this and other units are reported in the reviewed SAE papers. The emission factors were reproduced exactly in almost all the cases barring a few where the recalculated emission factors were off by just one significant digit. For example, in one case the reported emission factor was 65.0 g/h and the value calculated was 64.9 g/hp-h. The recalculation of emission factors served as a quality assurance step.

4.2 Comparing The Test Cycles

Emission factors were reported in the literature based upon a variety of test procedures. Before grouping data from multiple test procedures into a single database for statistical analysis, it is important to first identify whether the data from two or more different types of tests are comparable. Therefore, an initial task in data analysis was to identify each test procedure and to determine whether emissions from different test procedures could be converted to a consistent basis. The various cycles were compared on the basis of how their modes are defined. In all of the comparisons, the CARB J1088 test cycle was taken to be a standard with which all the others were compared. This is because most of the engine test data that are well documented used the CARB J1088 test procedure. Table 1 gives a brief description of the various test cycles.

Each test cycle is characterized by a number of steady-state "modes." Each steady-state mode typically involves operation at a specified engine speed or type of speed and load (i.e. idle, low, intermediate or rated) for a given length of time. The intermediate speed varies from one test cycle to another. For example, in the CARB J1088, C6M and 6-Mode test procedures the intermediate speed is 85 percent of the rated speed[18, 27] but for the EMA 13 mode test it is 60 percent of the rated speed[20]. The SAE J1088 test procedure has equal number of modes with intermediate speeds of 60 percent and 85 percent of the rated speed[21].

The CARB J1088 procedure proposes two test cycles for measuring emissions: one for engines used in hand held applications and the other for engines used in non-handheld applications. For handheld applications only the first two modes are used in testing, whereas for non-handheld applications modes 1, 3, 4, 5, 6, and 7 are used.[18]

There were a number of conclusions drawn from the comparison of test cycles. The EMA 13 Mode procedure is not the same as the CARB J1088. The EMA procedure has 9 steady state

Table 1. Comparison of Selected L&G Equipment Test Cycles Based on the Number of Modes and Type of Engine Load

No.	Cycle	No. of Modes	Number of Modes by Type of Engine Load		
			Low Idle	Intermediate	Rated
1	SAE J1088[21]	16	2	13	1
2	CARB J1088 (handheld)[18]	2	1	-	1
3	CARB J1088 (non-handheld)[18]	6	1	5	-
4	EMA 13-Mode[20]	13	3	5	5
5	C6M[27]	6	1	5	-
6	6-Mode[27]	6	1	5	-
7	Lawn Mower Cycle[24]	5	-	5	-
8	Generator Set Cycle[24]	5	-	-	5

modes that are not the same compared to any of the modes used in the CARB J1088 procedure. The C6M test procedure is the same as the CARB J1088 procedure for non-handheld engines because it uses the same steady state modes and the same weights. The Generator Set Cycle and the Lawn Mower Cycle use different steady state modes compared to the CARB J1088 procedure.[24] The Lawn Mower Cycle has modes with different loads and the Generator Set Cycle has modes with different speeds. The 6-Mode test cycle has the same modes as those defined in the CARB J1088 test cycle for non-handheld engines but the weightings used were not available. The engines which were tested using the 6 Mode test cycle were also tested using the C6M procedure. For such engines the data from C6M test procedure were used in the database. This was done because more information was available for the C6M procedure such as modal weightings. Also, the emission factors reported did not differ by more than six percent between the C6M and 6-Mode test cycles.

The SAE J1088 procedure from which the CARB J1088 procedure was derived has more modes than the CARB J1088 procedure. However not all of these are used to test a particular L&G engine. All steady state modes in the CARB J1088 procedure are in the SAE J1088, and it is mostly these modes that are recommended by the SAE J1088 procedure for testing L&G engines. For these reasons the emission factor data from L&G engines using these two test procedures were combined in the database.

4.3 Summary of the Database

In Table 2, a summary of the database is presented which shows the number of engines for which emission factor data could be obtained with respect to possible emission factor categories, pollutants, and emission factor units. The database includes data only from J1088 and C6M test procedures. Data for a total of 51 engines are included. The numbers shown for the 2-Stroke and the 4-Stroke category include both handheld and non-handheld engines. Please note that numbers in any given row or column do not add up to the total number of engines in the database because many engines can be included in more than one category (e.g., as 4-stroke, OHV, and Non-handheld) and may include data for more than one emission factor unit (e.g., g/hr, g/hp-hr, and g/gal).

Table 2. Summary of the L&G Engine Emission Factor Database

Category	Number of Engines					
	THC (g/hr)	**THC (g/hp-hr)**	**THC (g/gal)**	**NO_x (g/hr)**	**NO_x (g/hp-h)**	**NO_x (g/gal)**
2 – Stroke	18	16	11	18	16	11
4 – Stroke	19	22	12	19	27	12
OHV	9	9	4	9	10	4
LHV	6	9	6	6	13	6
Handheld	16	14	9	16	14	9
Non Handheld	21	24	14	21	29	14

Emission factor data were not available for all three emission factor units (i.e. g/hr, g/hp-hr, and g/gal) for all cases. For example, some reports had THC and NO_x data in g/hp-hr units only, while some had enough additional information provided so that emission factors could be calculated in other units. Also in one of the reports, instead of THC, the emission factors reported were in terms of Non-Methane Organic Gas (NMOG) and were not used in the database.[23] Thus, for the 4-stroke L&G category there are 27 data points for NO_x (g/hp-h) and only 22 data points for THC (g/hp-h).

Each data point in the database represents a separate make/model of L&G engine. However, there were three cases (2-stroke NO_x g/h, 2-stroke THC g/h and 4-stroke NO_x g/hp-h) when there were multiple emission factors reported for a particular engine measured by the same test procedure but under different test conditions or fuels. The value considered most representative of typical measurements was included in the database. In all of these cases the chosen value was not different from the other measurements by more than 11 percent.

4.4 Grouping the Test Data

The database compiled from the available literature was analyzed statistically to identify whether there are any significant differences in emissions that justify categorization of the data with respect to the type of engine application, the engine technology, and/or the engine size.

4.4.1 Categorizing Emission Factors By Application

The L&G data set compiled from available literature was evaluated to determine whether there is a statistically significant difference in the mean emission factor estimate when comparing engines used in handheld applications versus those used in non-handheld applications. For this, 2-tailed *t* – tests for the difference in means at a 5 percent level of significance were done on the data for these two categories. The comparison between handheld and non-handheld applications was made for six groups of data: for each of the two pollutants (NO_x and THC), three different emission factor units were considered separately (g/hr, g/hp-hr, and g/gallon). The *t* – Test results are described in Table 3. Except for NO_x emissions when measured in g/gallon units and THC emissions when measured in g/hp-h units, for all of the other cases considered there was a significant difference in the mean emission rates when comparing handheld and non-handheld datasets. Thus, it would be reasonable to separately calculate emission factors with respect to these two types of applications.

Table 3. Results of Comparison of Non-Handheld versus Handheld Emissions Data Based Upon *t*-Tests for the Difference of Two Means at a 5% Level of Significance.

Comparison	Pollutant	Units	Significant Difference between Data Sets?
Non-Handheld Engines Versus Handheld Engines	THC	g/h	Yes
		g/hp-h	No
		g/gallon	Yes
	NO_x	g/h	Yes
		g/hp-h	Yes
		g/gallon	No

Table 4. Results of Comparison of 2-Stroke versus 4-Stroke Engine Emissions Data, and of OHV versus LHV 4-Stroke Engines, Based Upon *t*-Tests for the Difference of Two Means at a 5% Level of Significance**.**

Comparison	Pollutant	Units	Significant Difference between Data Sets?
2-Stroke Engines versus 4-Stroke Engines	THC	g/h	Yes
		g/hp-h	Yes
		g/gallon	Yes
	NO_x	g/h	Yes
		g/hp-h	Yes
		g/gallon	Yes
4-Stroke Engines: OHV versus LHV	THC	g/h	Yes
		g/hp-h	No
		g/gallon	No
	NO_x	g/h	No
		g/hp-h	No
		g/gallon	No

4.4.2 Categorizing Emission Factors By Technology

The L&G data set compiled from available literature was evaluated to determine whether there is a statistically significant difference in the mean emission factor estimates when comparing 2-stroke versus 4-stroke engines. In addition, a comparison was made between OHV and LHV 4-stroke engines. To make the comparisons, 2-tailed *t* – tests for the difference in means at a five percent significance level were done. Both comparison were made for six groups of data: for each of the two pollutants (NO_x and THC), three different emission factor units were considered separately (g/hr, g/hp-hr, and g/gallon). The *t*-test results are described in Table 4. The analysis showed that there is a significant difference in the means of the 2-stroke versus 4-stroke engine data sets. Also except for THC emissions when measured in units of g/h, there was not a significant difference in the mean OHV versus LHV emissions. Thus, it is reasonable to develop emission factors separately for 2-stroke and 4-stroke engines. However, it is not necessary to subdivide 4-stroke engines into categories of OHV and LHV technologies.

4.4.3 Categorizing Emission Factors By Engine Size

Although in the NEVES report EPA has not categorized L&G engines of less than 25 hp by size, EPA plans to do so in its new NONROAD model. Within these size categories emission factors would be divided on the basis of application also if there were sufficient information to justify their use. Specifically, EPA reports that two stroke engines will be divided into a category of less than 3 hp and a category of greater than 3 hp. Both of these categories are assumed by EPA to correspond to non-handheld applications of the 2-Stroke engines. EPA does not report plans to categorize L&G four stroke engines by size.[31]

In the database, there are only 2 data points in the category of 2-stroke engines, non handheld (≤ 3hp) and only one data point in the category of 2-stroke engines, general non handheld (≥ 3hp). This is not sufficient data to perform a statistical analysis of uncertainty, nor is it a sound basis for subdividing the emissions database. Because 2-stroke engines are primarily used in handheld applications, and because 4-stroke engines are primarily used in non-handheld applications, there are few data available for non-handheld 2-stroke engines and for handheld 4-stroke engines. Therefore, on the basis of lack of data, the analysis of emissions with respect to size focuses on the type of engine, rather than on the type of application. For 2-stroke engines, the emissions measurements for the three non-handheld engines in the database fall within the range of measurements for the 16 handheld engines included in the database. Therefore, it was judged acceptable to combine the non-handheld 2-stroke engine data with the larger amount of handheld 2-stroke engine data to create a database with measurements for a total of 19 engines. Similarly, for 4-stroke engines, the emissions measurement for one handheld engine was within the range of values for 31 non-handheld 4-stroke engines. Therefore, it was judged acceptable to include the data for the handheld 4-stroke engine as part of the 4-stroke engine database of 32 total measurements.

To evaluate whether there is a possible dependence of emissions with respect to size for the 2-stroke and for the 4-stroke engines, scatter plots of emissions versus measures of size were created for each type of engine. Examples of these scatter plots are shown in Figures 1 and 2 for 4-stroke and 2-stroke engines, respectively. The example scatter plot for 4-stroke engine THC emissions versus engine horsepower indicate a possible difference in emission rates for engines of less than approximately eight horsepower compared to those in the 8 to 25 hp range. The differences are two-fold: (1) there appears to be substantially more variability in emissions for the smaller size range, with emissions varying from approximately 10 g/hp-hr to as much as nearly 100 g/hp-hr, versus values close to 10 g/hp-hr in the upper size range; and (2) the average emissions appear to be greater for the smaller size range.

The example scatter plot for 2-stroke engine THC emissions versus engine rated horsepower does not provide any strong indication of a dependence of emissions with respect to engine size. There is more variability in the emissions for engines of less than 3 hp compared to engines of greater than 3 hp. However, the average values of emissions appear to be approximately similar for both size ranges. Therefore, it is less clear in this case as to whether engine emissions differ significantly with respect to size.

Figure 1. Scatter Plot of THC (g/hp-h) Emission Factors for 4-Stroke L&G Engines Versus the Size of the Engine (Rated Horsepower).

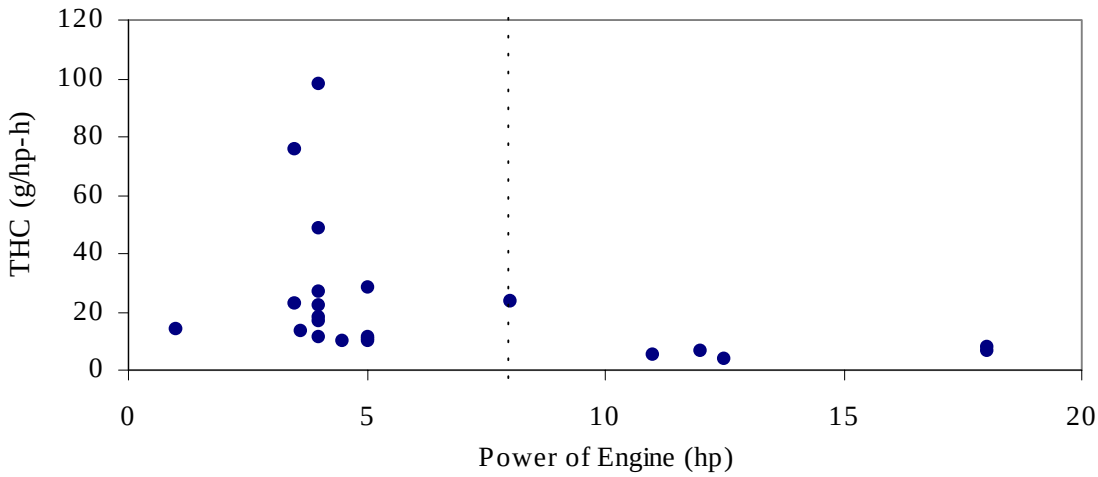
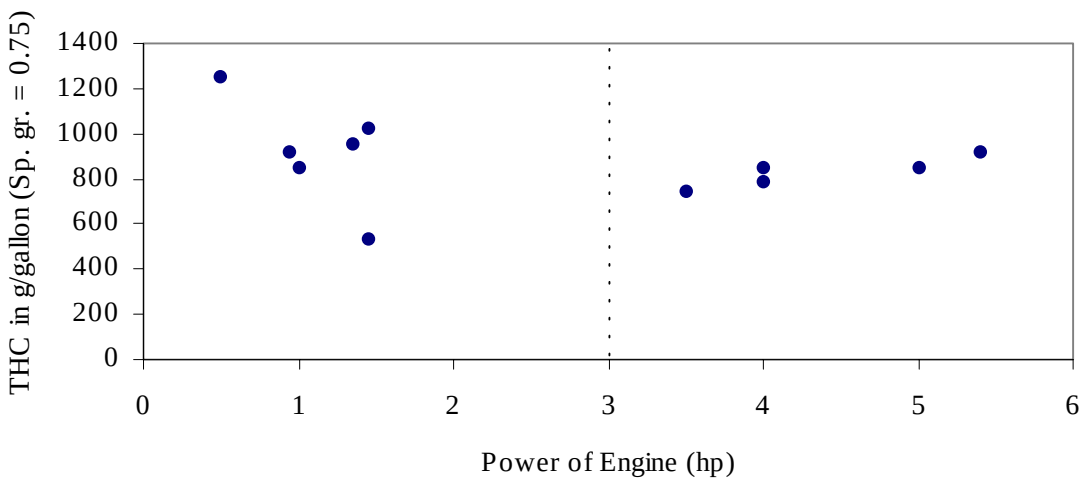


Figure 2. Scatter Plot of THC (g/gallon) Emission Factors for 2-Stroke L&G Engines Versus the Size of the Engine (Rated Horsepower).



To identify whether apparent differences in the scatter plots are statistically significant, two-tailed t-tests, at a five percent significance level, for differences in the mean values of emissions for different size ranges were conducted. For the 4-stroke engines, emissions data were divided into two size ranges of less than eight horsepower and greater than eight horsepower. Six comparisons were made inclusive of both pollutants and of all three units used for emission factors. Similarly, for 2-stroke engines, six pair-wise comparisons were made for engines sizes of less than three horsepower versus those greater than three horsepower. The results of the statistical tests are given in Table 5.

The statistical test results depend on which emission factor unit is used. For example, for 2-stroke engines there is not a statistically significant difference in mean values if g/gallon emission factors are used. For 2-stroke engines, the 95 percent confidence interval for the mean of the small size range is 730 g/gallon to 1,100 g/gallon, compared to a 95 percent confidence

Table 5. Results of Comparison of Size Ranges for 4-Stroke and for 2-Stroke Engines Based Upon *t*-Tests for the Difference in Mean Emissions at a 5% Level of Significance**.**

Comparison	Pollutant	Units	Significant Difference between Data Sets?
4 Stroke Engines: Size Range ≥ 8hp vs. Size Range < 8hp	THC	g/h	No
		g/hp-h	Yes
		g/gallon	Yes
	NO_x	g/h	Yes
		g/hp-h	No
		g/gallon	Yes
2 Stroke Engines: Size Range ≥ 3hp vs. Size Range < 3hp	THC	g/h	Yes
		g/hp-h	No
		g/gallon	No
	NO_x	g/h	No
		g/hp-h	No
		g/gallon	No

interval for the larger size range of 770 g/gallon to 890 g/gallon. The confidence interval for the larger size range is enclosed by the confidence interval for the smaller size range, indicating substantial overlap between the two intervals. As indicated in Table 5, the mean emissions for these two size ranges are not statistically significantly different from each other.

Based upon the results in Table 5, it appears that, for 2-stroke engines, there is little evidence to suggest that the mean emissions are significantly different between the two size ranges considered. The only exception to this is for g/hr units for the NO_x emission factor. However, because g/hr emissions are not normalized to any measure of engine size (e.g., hp-hr produced or gallons of fuel consumed) it is expected that this measure of emissions would be size dependent. Furthermore, g/hr emission factors are not particularly useful as an emission factor because they are size dependent. Thus, for practical purposes, there is not a demonstrated need, based upon the data, to divide 2-stroke engines into size categories. For 4-stroke engines, there is more evidence that there are differences in emissions for different size ranges. For example, average g/gallon emissions differ significantly between the two size ranges. However, as noted, mean g/hp-hr emissions do not differ significantly.

The findings from comparison of different possible categories of data are as follows: (1) there are significant differences in mean emissions between 2-stroke engines and 4-stroke engines; (2) there are significant differences in mean emissions between handheld and non-handheld engines; (3) there are not significant differences in emissions between 4-stroke OHV engines and 4-stroke LHV engines; and (4) there are mixed indications regarding differences in mean emissions with respect to engine size. Because there is strong concordance between 2-stroke engines and handheld engines, and between 4-stroke engines and non handheld engines, and lack of sufficient data for 2-stroke non-handheld engines and 4-stroke handheld engines, classifications with respect to engine design (2 versus 4 strokes) or application (handheld versus non-handheld) are approximately equivalent. In this work, classification by engine design is selected. The analyses

suggest no need to classify 4-stroke engines with respect to OHV versus LHV design. There is not a strong basis for classifying either engine design with respect to size ranges, especially for g/hp-hr emission factors. Therefore, the only classification scheme selected is with respect to engine design.

5.0 ANALYSIS OF VARIABILITY AND UNCERTAINTY

The previous sections focused upon development and exploration of an L&G engine emission factor database. In this section, variability in the emission factor data are quantified using parametric probability distribution models. Uncertainty in the mean emission factors, and regarding the fitted distributions, is quantified using bootstrap simulation. Estimates of uncertainty in mean emission factors are presented for both 2-stroke and 4-stroke engines for two pollutants (THC and NO_x) and for three emission factor units (g/hr, g/hp-hr, and g/gallon).

5.1 Quantification of Inter-Engine Variability in Emissions

Measured emissions vary from one engine to another in the database. The variability in emissions can be described as an empirical cumulative distribution function (CDF) or by a parametric probability distribution function. Parametric distributions offer advantages of enabling interpolation within the range of the observed data and for extrapolation to the tails of the distribution beyond the range of observed data.[10] The latter is important because it is unlikely, with a finite sample of data, that the true minimum or maximum values are represented by the sample minimum and maximum values. In this work, alternative parametric probability distribution models were fit to the data and evaluated for goodness-of-fit based upon visualization of the fitted distribution compared with the data and based upon the results of bootstrap simulations of the fitted distribution. Statistical goodness-of-fit tests were not used because the data sample sizes are too small. The parametric distributions considered were Normal, Lognormal, Gamma and Weibull. Maximum likelihood estimation was used to estimate the parameters of the fitted distributions. Evaluation of the adequacy of fit is further addressed in the following sections.

5.2 Quantification of Uncertainty in Mean Emissions and Fitted Distributions Using Bootstrap Simulation

In order to characterize uncertainty in the mean emission factors and to evaluate the adequacy of the fit of a parametric distribution to the data, bootstrap simulation was performed. Bootstrap simulation is a numerical technique originally developed for the purpose of estimating confidence intervals for statistics.[32] This method can provide solutions for confidence intervals in situations where exact analytical solutions may be unavailable and in which approximate analytical solutions are inadequate. Example applications of bootstrap simulation are available elsewhere.[14,16,33,34]

Bootstrap simulation uses a conceptually straightforward approach. By fitting a parametric distribution to a dataset, an estimated population distribution is developed. A simulated random sample of data points of the same sample size as the original data is drawn from the assumed population distribution. This data set is referred to as a “bootstrap sample.” Any statistics of interest, such as the mean, standard deviation, distribution parameters, distribution percentiles, or others, are calculated from the bootstrap sample. For each bootstrap sample, the estimate of a

statistic is referred to as the "bootstrap replication" of the statistic. The process is then repeated many (e.g. 500 to 2000) times to create a probability distribution of bootstrap replications of a statistic. A probability distribution for a statistic is referred to as a "sampling distribution." Confidence intervals for a statistic are inferred from its sampling distribution. For example, the 2.5^{th} and 97.5^{th} percentiles of the sampling distribution would enclose a 95% confidence interval. Using bootstrap simulation, confidence intervals are constructed for the mean emission factor estimates and for the fitted cumulative distribution functions.

5.3 Results of Quantitative Analysis of Variability and Uncertainty in L&G Engine Emission Factors

The probabilistic techniques described in the previous two sections were applied to each of the emission factor data groups identified as having statistically significant differences in mean emissions with respect to each other. To illustrate the application of the probabilistic methods, example results are presented for selected cases.

In Figure 3, results are shown for the 2-stroke engine NO_x emission data set in units of g/hp-hr. The variability in emissions among these engines is from approximately 0.1 g/hp-hr to approximately 2.4 g/hp-hr, with most of the data having emission values of less than 1.0 g/hp-hr. A fitted Lognormal distribution is shown in comparison to the data. The Lognormal distribution captures the overall trends of the empirical distribution of the data. There is some scatter of the data above and below the fitted distribution. In order to evaluate whether the deviations of the data with respect to the fitted distribution imply a poor fit, confidence intervals for the fitted CDF of the Lognormal distribution are shown. These confidence intervals were obtained based upon bootstrap simulation. Approximately one half of the data (7 out of 16 data points) are enclosed by the 50 percent confidence interval, and nearly all of the data are enclosed by the 95 percent confidence interval. This comparison suggests that the Lognormal distribution is an adequate fit to the data set and, therefore, is a reasonable representation of inter-engine variability in emissions. Similar results are shown for the same engine category with respect to THC emissions in Figure 4. In this latter case, it is clear that the fitted distribution is heavily influenced by an extreme data point with an emission rate of approximately 1,300 g/hp-hr. In contrast, all of the other data have emission rates of approximately 300 g/hp-hr or less. However, we have no basis for discarding the extreme data point. Furthermore, if this large data point were ignored, the mean emission estimate would be biased to a lower value than if this data point is retained in the analysis. Although there appear to be some systematic deviations between the data and the fitted distribution in a cumulative probability range between approximately 0.6 and 0.9, almost all data are enclosed by the 90 percent confidence interval of the bootstrap simulation. Therefore, the fit appears to be adequate even though it is not perfect.

Figures 5 and 6 illustrate probabilistic analysis results for 4-stroke engines for NO_x and THC emission factors, respectively, in units of g/hp-hr. The inter-engine variability in NO_x emissions among the 27 data points ranges from approximately 0.4 g/hp-hr to over 5.0 g/hp-hr. There are no obvious outliers in the data set; the highest value differs from the second highest value by only approximately 35 percent. The fitted Lognormal distribution agrees well with the lower and upper portions of the distribution, although there is apparent disagreement near the central values of the distribution. However, all of the data are enclosed within the 95 percent confidence interval, indicating that the fit is adequate.

Figure 3. Fitted Lognormal Distribution and Bootstrap Simulation Results for 2-Stroke L&G Engine NO_x Emission Data in Units of g/hp-hr.

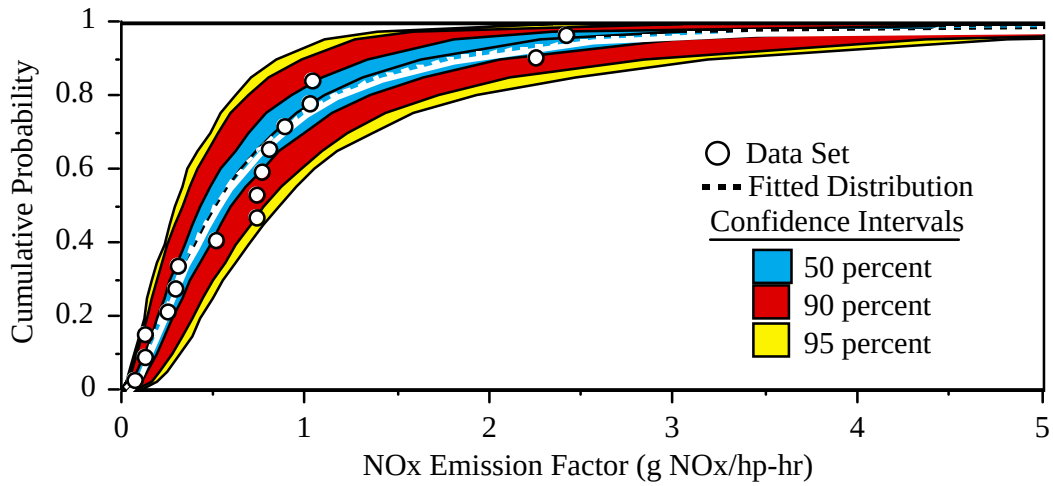
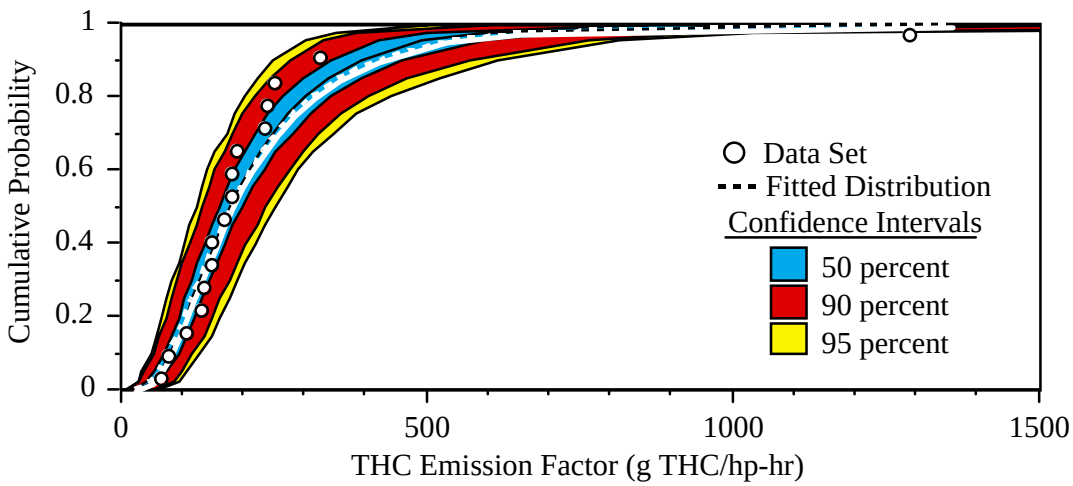


Figure 4. Fitted Lognormal Distribution and Bootstrap Simulation Results for 2-Stroke L&G Engine THC Emission Data in Units of g/hp-hr.



In the case of the THC data, the inter-engine variability in emission ranges from approximately 5 g/hp-hr to nearly 100 g/hp-hr. There are three data points above 30 g/hp-hr. The fitted Lognormal distribution appears to slightly overestimate the emission values systematically in a range of cumulative probability from 0.3 to 0.9. However, all of the data in this range are enclosed by the 90 percent confidence interval. Thus, the fit appears to be adequate.

Results of the analysis of uncertainty in mean emission factors for all cases considered are given in Table 6. For example, for the 2-stroke engine NO_x emissions in units of g/hp-hr depicted in Figure 3, the mean emission rate is 0.81 g/hp-hr. This corresponds to the approximately 65th percentile of the fitted distribution. The mean occurs at a cumulative probability above the median (50th percentile) because the data and the distribution are positively skewed. The 95 percent confidence interval for the mean is 0.44 g/hp-hr to 1.34 g/hp-hr, which is a range of minus 46 percent to plus 65 percent of the mean value. The relatively wide range of uncertainty in the mean emission factor in this case is a result of both the relatively small sample size of this data set (n=16) and the wide range of variability in the data.

Figure 5. Fitted Lognormal Distribution and Bootstrap Simulation Results for 4-Stroke L&G Engine NO_x Emission Data in Units of g/hp-hr.

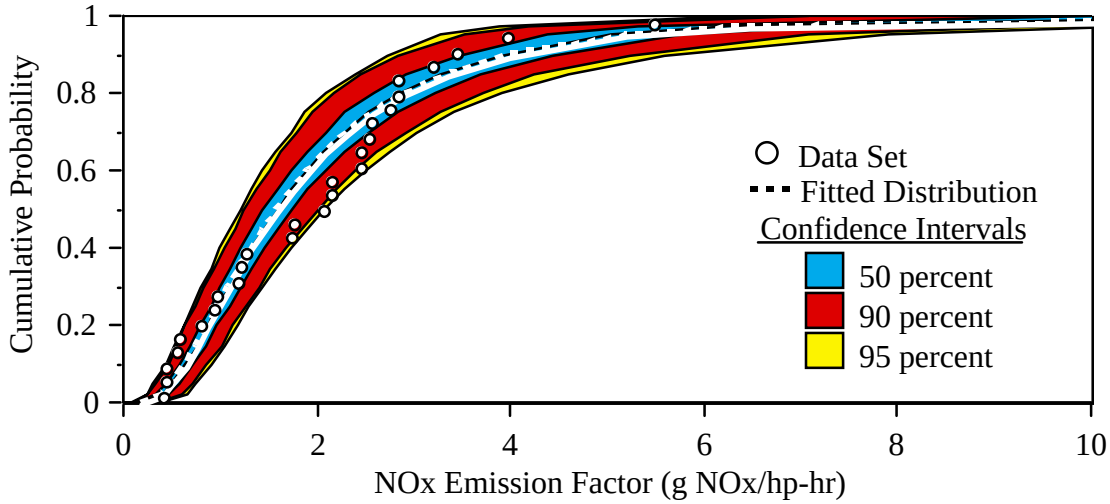
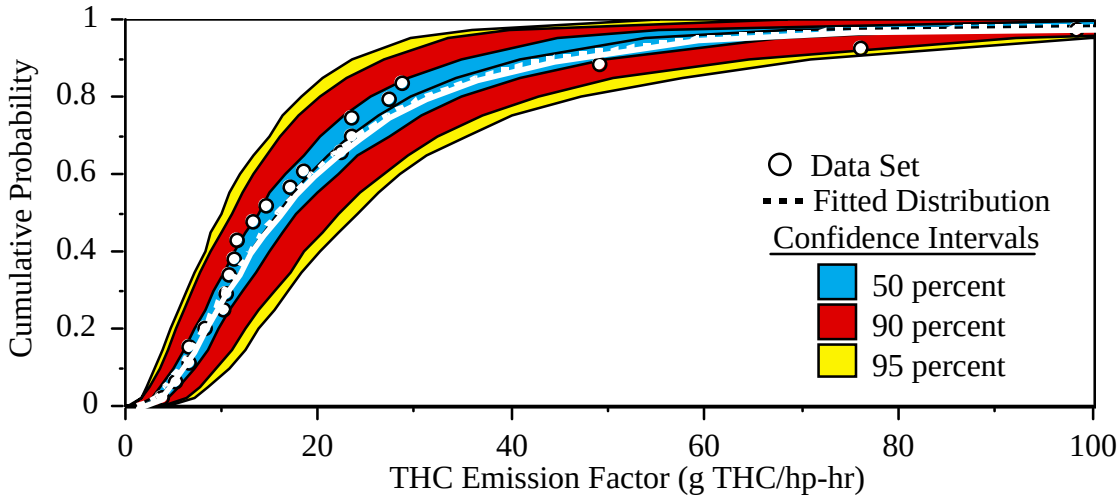


Figure 6. Fitted Lognormal Distribution and Bootstrap Simulation Results for 4-Stroke L&G Engine THC Emission Data in Units of g/hp-hr.



Of the 12 cases shown in Table 6, all but one have uncertainty ranges of greater than approximately plus or minus 20 percent, and six have uncertainty ranges of greater than approximately plus or minus 40 percent. Thus, there is substantial quantified uncertainty in the mean emission factors. For the same type of engine and pollutant, the range of uncertainty in emission factors varies depending on the unit used, at least in part because differing numbers of data points are available depending on the unit of measure employed.

Table 6. Results of the Uncertainty Analysis of Mean NO_x and THC Emission Rates for 2-Stroke and 4-Stroke Lawn and Garden Engines for Three Emission Factor Units (g/gal, g/h, and g/hp-h)

Engine Type	Pollutant	Units	Number Of Data	Fitted Dist.	Mean	95% C.I. on Mean[a]	Relative Uncertainty[b]
2-Stroke	NO_x	g/gal	11	Gamma	4.38	2.4 - 6.87	-45% to +57%
		g/h	18	Lognormal	0.92	0.53 - 1.43	-42% to +55%
		g/hp-h	16	Lognormal	0.81	0.44 - 1.34	-46% to +65%
	THC	g/gal	11	Lognormal	877	769 – 984	-12% to +12%
		g/h	18	Lognormal	235	175 – 303	-26% to +29%
		g/hp-h	16	Lognormal	222	152 – 306	-32% to +38%
4-Stroke	NO_x	g/gal	12	Weibull	13.2	8.43 – 18.2	-36% to +38%
		g/h	19	Lognormal	5.48	2.69 – 10	-51% to +82%
		g/hp-h	27	Lognormal	2.05	1.54 - 2.82	-25% to +38%
	THC	g/gal	12	Lognormal	124	82.4 – 186	-34% to +50%
		g/h	19	Lognormal	34.2	24.8 – 47.2	-27% to +38%
		g/hp-h	22	Lognormal	21.5	13.3 – 31.1	-38% to +45%

[a] Numbers shown here are for the 95 percent confidence interval of the mean obtained from bootstrap simulation in terms of absolute emission rates

[b] Numbers shown here are for the 95 percent confidence interval of the mean obtained from bootstrap simulation in terms of relative deviation from the mean.

6.0 CONCLUSIONS AND RECOMMENDATIONS

This paper has successfully demonstrated a procedure for quantifying variability in emissions and uncertainty in mean emissions using parametric probability distributions and bootstrap simulation. The procedure has been demonstrated with respect to a case study of lawn and garden equipment of less than 25 hp.

A key difficulty encountered in this work was to obtain a well-documented and complete database. Regulatory agencies that make public policy should, as a matter of routine, make publicly available the data used in calculations used to develop emission factors. Such data should be made widely available through technical documents and databases on the World Wide Web. While in some cases data may be of proprietary origin, information that would identify a specific emission source with a particular company or facility need not be reported. Furthermore, in reporting values of emission factors, quantitative information regarding the variability in emissions and regarding the uncertainty in the average emissions should be provided.

In this paper, uncertainty associated with statistical random sampling error was quantified. However, uncertainty associated with lack of representativeness and with measurement errors was not quantified. Each of these points is further addressed here.

Measurements of L&G engines have been made using a variety of test procedures. These procedures are not always comparable with each other. Care was taken in this work not to group together data that were obtained from dissimilar test methods. For emission factor purposes, it is

important to have data that are representative of real world applications of these types of engines. Therefore, candidate test methods should be evaluated with respect to their representativeness. In this paper, we have not attempted to quantify uncertainty associated with potential lack of representativeness. Such an analysis would require at least some benchmark measurements of real-world in-use emissions, which is a recommendation to regulatory agencies for future work.

Uncertainty in emissions measurements can, potentially, be a significant source of uncertainty, particularly for some pollutants and/or emission sources. In the judgment of the authors, the methods used to measure engine emissions are relatively well-known and of reasonably high quality. Therefore, it is expected that the measurement errors are not large with respect to the quantified uncertainty associated with random sampling error, except in cases when this quantified uncertainty is relatively small (e.g, plus or minus 10 percent). However, verification of this assumption is a need for future work. An obstacle to verifying this assumption is the lack of information regarding the precision and accuracy of the test methods. Such information should be provided in any emission test report.

In this work, decisions regarding how to categorize the emission factor database were made based upon empirical evidence from the database regarding whether mean emissions differed in a statistically significant manner for different subgroups. Significant differences were found between 2-stroke and 4-stroke engines for both NO_x and THC, and also between handheld and non-handheld applications. Because handheld applications typically involve 2-stroke engines, and nonhandheld applications typically involve 4-stroke engines, it is not useful to use both sets of categories. Therefore, only the first one was chosen. There was no significant difference in the mean emissions of 4-stroke OHV versus LHV engines. There is some evidence to suggest that emissions may differ by size for both 2-stroke and 4-stroke engines, but the evidence was judged not to be sufficiently compelling at this time.

Inter-engine variability in emissions was found to be substantial, such as a factor of 10 or more variation from the smallest to largest values in a given data set. Although it is clear that there are often only a small number of large values in a given data set, unless there are known errors in the data, it is not appropriate to discard such values. Because of the relatively small data set sample sizes and the large range of variability, the uncertainty in the mean emissions were relatively large, with most cases evaluated having uncertainty ranges for the mean in excess of plus or minus 20 percent.

The large range of quantified uncertainty in emission factors suggests that it is important to quantify uncertainty. As the National Research Council noted in its recent report on modeling mobile source emissions, it is not possible to quantify all sources of uncertainty. Nonetheless, the quantifiable portion of uncertainty should be taken into account when reporting and using emission factors. Non-quantifiable contributions to uncertainty should be acknowledged qualitatively. Decision-makers should be aware of both the strengths and limitations of emission factors and emission inventories, so that decisions regarding air quality management can be made that are robust to uncertainty. Furthermore, by understanding the key sources of uncertainty in an emission inventory, resources can be prioritized to reduce uncertainty, such as by collecting better and more data. Thus, the probabilistic methodology presented here is part of an overall approach to developing policy, program management, and research planning.

7.0 ACKNOWLEDGMENTS

This work was supported by U.S. Environmental Protection Agency Science to Achieve Results (STAR) Grants Nos. R826766 and R826790. Although the research described in this paper has been funded wholly or in part by the U.S. EPA's STAR grant program, it 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.

The authors are grateful to Mr. James Warila of the U.S. EPA's Office of Mobile Sources in Ann Arbor, MI, for his assistance and recommendations regarding data sources. We also thank Mr. Charles T. Hare, Director Department of Emissions Research, Southwest Research Institute, for kindly providing us with emissions test reports.

8.0 REFERENCES

1. National Research Council. Modeling Mobile-Source Emissions. National Academy Press: Washington, DC. 2000.
2. US Environmental Protection Agency. *Reducing Air Pollution From Nonroad Engines*; EPA420-F-00-048, at EPA-OTAQ Nonroad Engine Emissions Page. http://www.epa.gov/otaq/regs/nonroad/f00048.htm (accessed January 2001)
3. US Environmental Protection Agency. *Compilation of Air Pollutant Emission Factors, AP-42, Fifth Edition*; US Environmental Protection Agency, 1995
4. Beck, L.; Wilson, D. "EPA's Data Attribute Rating System." In *Emission Inventory: Planning for the Future, The proceedings of A Specialty Conference*, Air & Waste Management Association: Pittsburgh, PA 1997: pp 176-189.
5. Dickson, R.J., Hobbs A. D.; "Evaluation of Emission Inventory Uncertainty Estimation Procedures." Paper No. 89-24.8, In *82nd Annual Meeting*, Air & Waste Management Association: Anaheim, CA 1989.
6. National Research Council. *Rethinking the Ozone Problem in Urban and Regional Air Pollution*, National Academy Press: Washington, DC, 1991.
7. Balentine, H. W.; Dickson, R. J. "Development of Uncertainty Estimates for the Grand Canyon Visibility Transport Commission Emission Inventory." In *Conference on the Emission Inventory*, Air & Waste Management Association: Research Triangle Park, NC, 1995.
8. National Research Council; *Science and Judgement in Risk Assessment*, National Academy Press: Washington, DC, 1994.
9. US Environmental Protection Agency. *Guiding Principles for Monte Carlo Analysis*; EPA/630R-97/001, US Environmental Protection Agency, 1997.
10. 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.
11. Kini, M.D., and H.C. Frey, Probabilistic Evaluation of Mobile Source Air Pollution: Volume 1, Probabilistic Modeling of Exhaust Emissions from Light Duty Gasoline Vehicles, Prepared by North Carolina State University for Center for Transportation and the Environment: Raleigh, December 1997.
12. Pollack, A.K., P. Bhave, J. Heiken, K. Lee, S. Shepard, C. Tran, G. Yarwood, R.F. Sawyer, and B.A. Joy, "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.

13. Frey, H.C., R. Bharvirkar, and J. Zheng, Quantitative Analysis of Variability and Uncertainty in Emissions Estimation, Prepared by North Carolina State University for the U.S. Environmental Protection Agency, Research Triangle Park, NC. July 1999. H. C. Frey Home Page. http://www4.ncsu.edu/~frey/freytech.html (accessed January 2001).
14. 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.
15. Rhodes, D.S., and H.C. Frey, "Quantification of Variability and Uncertainty in AP-42 Emission Factors Using Bootstrap Simulation," Emission Inventory: Planning for the Future (held October 28-30 in Research Triangle Park, NC), Air and Waste Management Association: Pittsburgh, Pennsylvania, October 1997, pp. 147-161.
16. Frey, H.C.; Li, S., " Quantification of Variability and Uncertainty in Natural Gas-fueled Internal Combustion Engine NO_x and Total Organic Compounds Emission Factors," Proceedings, Annual Meeting of the Air & Waste Management Association: Pittsburgh, PA. June 2001 (in press).
17. California Air Resources Board. *Technical Support Document for California Exhaust Emission Standards and Test Procedures for 1994 and Subsequent Model Year Utility and Lawn and Garden Equipment Engines*; Attachment C to CARB Mail Out #90-64: El Monte, CA, 1990.
18. Carroll, J. N., *Emission Tests of In-use Small Utility Engines*; SwRI 3426-006, Task III Report, Nonroad Source Emission Factors Improvement, Prepared by Southwest Research Institute for U.S. EPA: Ann Arbor, MI. 1991.
19. Carroll, J. N.; White, J. J., *Nonroad Source Emission Testing of Small Utility Engines*, Prepared by Southwest Research Institute for Sierra Research: Sacramento, CA. 1993; SwRI 08-4855-220.
20. Hare, C. T.; Springer, K. J. *Exhaust Emissions from Uncontrolled Vehicles and Related Equipment Using Internal Combustion Engine;* Final report, Part 4, Small Air-cooled Spark Ignition Utility Engines, Contract EHS 70-108, Prepared for Emission Characterization and Control Development Branch, Office of Mobile Source Air Pollution Control and National Air Data Branch, Office of Air Quality Planning and Standards, Office of Air and Water Programs, U.S. EPA.
21. White, J. J., *Toro Small Air-cooled Engine Exhaust Emissions Baseline*; Appendix A, D, E, G and H, prepared by Southwest Research Institute for The Toro Company: South Minneapolis, MN.1990; SwRI-08-3152.
22. Carroll, J. N.; White, J. J., *Emission Testing of In-use Handheld Engines;* SwRI 4855-211, Prepared by Southwest Research Institute for Sierra Research, Sacramento: CA, 1994.
23. Gabele, P." Exhaust Emissions from Four-Stroke Lawn Mower Engines," *J. Air & Waste Manage. Assoc*., 1997, 47, pp. 945-952.
24. Sun, X.; Brereton, G.; Morrison, K.,; Patterson, D., "Emissions Analysis of Small Utility Engines," Society of Automotive Engineers Technical Paper No. 952080, 1995.
25. Hare, C. T.; Springer, K. J.; Oliver, W. R.; Houtman, W. H., " Small Engine Emissions and their Impact," Society of Automotive Engineers paper 730859, 1973.

26. Carroll, J. N.; White, J. J.; Hare, C. T.; Lourenco, G. J., “Emission Factors for Small Utility Engines,” Society of Automotive Engineers Technical Paper No. 910560, 1991.
27. Gabele, P., “Emissions from 4-Cycle Walk-Behind Mower Engines: Test Cycle Effects,” Society of Automotive Engineers Technical Paper No. 972793, 1997.
28. Carroll, J. N.; White, J. J.; Hare, C. T.; Lourenco, G. J., “Emission Control Strategies for Small Utility Engines,” SAE Technical Paper No. 911807, 1991.
29. Frey, H.C.; Eichenberger, D.A., *Remote Sensing of Mobile Source Air Pollutant Emissions: Variability and Uncertainty in On-Road Emissions Estimates of Carbon Monoxide and Hydrocarbons for School and Transit Buses;* FHWY/NC/97-005, pp 55, prepared by Department of Civil Engineering, North Carolina State University for North Carolina Department of Transportation, 1997.
30. US Environmental Protection Agency. *Nonroad Engine and Vehicle Emission Study (NEVES)*; EPA 460/3-91-02, NTIS Accession No. PB92-126960, Office of Mobile Sources, Certification Division: Ann Arbor, MI, 1991.
31. Beardsley, M.; Lindhjem, C. E.; Harvey, C., *Exhaust Emission Factors for Nonroad Engine Modeling - Spark Ignition*, Report No. NR-010b, U.S. EPA Office of Mobile Sources, Assessment and Modeling Division: Ann Arbor, MI, 1998, at EPA-OTAQ Nonroad Spark Ignition Engine Emissions Home Page. http://www.epa.gov/otaq/equip-ld.htm (accessed March 1999)
32. Efron, B.; Tibshirani, R. J., *An Introduction to Bootstrap*; Chapman and Hall: New York 1993.
33. Frey, H. C.; Burmaster, D. E., “Methods for Characterizing Variability and Uncertainty: Comparison of bootstrap Simulation and Likelihood-Based Approaches”; Risk Anal., 1999, 19, 109-129
34. Zheng, J.; Frey, H.C., "Quantitative Analysis of Variability and Uncertainty: An Illustration of Methods Using Mixture Distributions," Proceedings, Annual Meeting of the Air & Waste Management Association, Pittsburgh, PA, June 2001 (in press).

KEYWORDS
Uncertainty, Variability, Emission Factors, Lawn and Garden equipment, nonroad, nitrogen oxides, total hydrocarbons, data quality, bootstrap simulation