Quantification of Variability and Uncertainty for Selected Nonroad Mobile Source Emission Factors

H. Christopher Frey and Sachin Bammi
Department of Civil Engineering
Campus Box 7908
North Carolina State University
Raleigh, NC 27695-7908
frey@eos.ncsu.edu

ABSTRACT

Emission factors 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. Uncertainty refers to lack of knowledge regarding the true value of emissions. 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 (L&G) engines and for construction, farm, and industrial (CFI) equipment. Data were obtained from emissions testing conducted by others. Databases were created and statistically analyzed to determine the minimum number of emission source categories with statistically significantly different average emissions. Inter-engine variability in emissions was quantified using parametric distributions. Bootstrap simulation was used to characterize confidence intervals for the fitted distributions. For 2-stroke L&G engines, the 95 percent confidence intervals for the mean emission factors for total hydrocarbon (THC) and Nitrogen Oxides (NO_x) emissions in g/hp-hr units were -32% to +38% and -46% to +65%, respectively. For 4-stroke L&G engines, the confidence intervals for mean emissions in g/hp-hr units were -38% to +45% for THC and -25% to +38% for NO_x. For CFI engines, which are primarily diesel, the 95 % confidence intervals for the mean emission factors were as small as –10% to +11% and as large as –48% to +49%. 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 are discussed.

INTRODUCTION

The objective of this paper is to demonstrate a quantitative approach for quantifying variability and uncertainty in emission factors with application to nonroad mobile source emissions. This paper is part of a larger program at NC State University to develop and demonstrate methods for quantification of variability and uncertainty, apply the methods to emission factors for a variety of emission source categories, develop probabilistic emission inventories, and evaluate the policy implications of a probabilistic approach to development and use of emission inventories. As a case example, much of the work is aimed at developing a probabilistic emission inventory for use in tropospheric ozone air quality modeling. Therefore, the focus is on characterizing uncertainty in emissions of NO_x and hydrocarbons. In addition to the non-road emissions source categories addressed in this paper, other work at NC State focuses on the following source categories: highway vehicles; power plants; stationary natural gas-fired engines; gasoline terminals; consumer solvents; architectural coatings; and asphalt paving, among others. In addition to work focused on ozone precursors, work is also underway to develop and apply probabilistic methods to urban air toxic emission inventories.

THE NEED FOR PROBABILISTIC ANALYSIS

The National Research Council (NRC) and the U.S. Environmental Protection Agency (EPA) have increasingly recognized the need for a quantitative approach to dealing with uncertainty in environmental modeling applications. For example, the NRC recently issued a report regarding mobile source emissions that recommends that uncertainty in such emission be quantified.[1] In 1997 EPA issued a report entitled "Guiding Principles for Monte Carlo Analysis" that was motivated primarily by the need to quantify variability and uncertainty in the context of human exposure and risk analysis.[2] However, based upon the recommendation of the NRC, EPA will have to respond with methods for quantifying uncertainty for mobile source emissions.

Variability refers to real differences in emissions, such as from one engine to another. Uncertainty refers to lack of knowledge regarding the true value of emissions. Sources of uncertainty can include: random sampling error; measurement error; non-representative data; data entry errors; and/or lack of data. This paper focuses primarily upon random sampling error as a key source of uncertainty. As part of separate work now underway, NC State is exploring methods for incorporating knowledge regarding measurement error into uncertainty analysis. The issue of non-representative data is difficult to quantify in the absence of benchmark data regarding the true real-world in-use emissions of a particular source category. In the case of nonroad vehicles, the available emission measurements are typically made using test cycles involving a weighted average of measurements made in the laboratory over multiple steady state operating modes. These cycles may or may not produce emissions similar to those produced by the same equipment when operated in the field. The potential lack of representativeness of available laboratory data for nonroad emission sources could lead to biases in the emission factor estimates. In the future, it is expected that real-world in-use emissions data will be obtained from nonroad equipment using portable on-board instruments that are now becoming available. As part of a recent project, NC State has analyzed some examples of on-board data obtained from a limited set of diesel nonroad equipment, as well as for onroad vehicles.[3,4] However, the on-board nonroad data are not sufficient for inclusion in this analysis at this time.

In developing an emission inventory, the objective is typically to estimate average emission factors for a given source category to support a prediction of the total emissions in a given geographic area over some time period. Estimation of the uncertainty in the average emission factor is therefore the main focus of this paper. However, in order to estimate uncertainty associated with random sampling error, it is necessary to quantify the inter-engine variability in emissions and to employ appropriate statistical methods for inferring a probability distribution for uncertainty in the mean emissions.

The importance of a quantitative approach to uncertainty analysis has been highlighted by the NRC.[1] A comprehensive uncertainty analysis can be used to to assess the relative contributions to uncertainty in an emission inventory of individual emission source categories and of specific emission factors and/or activity factors. Such information can be used to identify the emission inventory inputs that contribute the most to uncertainty in total emissions. With this knowledge, data collection can be specifically targeted to obtain more or better data (e.g., real-world in-use data) to reduce uncertainty and improve the overall inventory. Thus, it is possible to develop a rational basis for guiding future emission testing programs and emission inventory improvement activities. Information about uncertainty is also important in conveying the quality of the emission factors, activity factors, and the emission inventory to enable decision-makers to account for uncertainty when making air quality management decisions. For example, when comparing emission estimates to an emissions budget for conformity or compliance purposes, a decision maker may wish to know the confidence with which an emissions goal can be met. Similarly, when making predictions with air quality models for comparison to ambient monitoring data for purposes of validation, the range of uncertainty in the air quality model prediction should be taken into account when determining whether there is agreement with monitoring data.

As the NRC indicates, it is difficult to quantify the uncertainty associated with non-representative samples.[1] "That a perfect assessment of uncertainty cannot be done, however, should not stop researchers from estimating the uncertainties that can be addressed quantitatively." The NRC recommends that EPA archive data used in model development and document analysis approaches. The NRC also recommends that future versions of the MOBILE model and other models should be developed to facilitate uncertainty analysis.

In the case of mobile sources, there is not an established practice for conveying the quality of the emission factors. However, in the case of stationary source emission factors, EPA has defined and reported qualitative “A” through “E” ratings.[5] 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.[6] While DARS can be used to compare quality ratings for EIs, it cannot be used to quantify the precision of an inventory or 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.[7,8,9] These approaches have various shortcomings. In many cases, assumptions are made that the probability distribution model representing uncertainty in an emission or activity factor is normal or lognormal without empirical justification. No distinction is made between inter-unit variability and uncertainty in estimating uncertainty in emissions and activity factors. Because the range of inter-unit variability is larger than the range of uncertainty in the average, an uncertainty analysis that is improperly based upon inter-unit variability will lead to an overestimation of uncertainty in the emission inventory.

The probabilistic approach presented addresses many of the concerns raised in the NRC and earlier studies. Specifically, the approach presented here properly distinguishes between inter-unit variability and uncertainty in the average; can be based upon a variety of assumptions regarding the shape of the estimated population distribution for inter-unit variability and need not be limited to normality or lognormality assumptions; is based upon analysis of empirical data; and produces distributions of uncertainty for the average emission factor that may take on an appropriate shape based upon the sample size of empirical data and the skewness of the estimated population distribution for inter-unit variability.

NONROAD EMISSION SOURCES AND DATA

The 1990 Clean Air Act Amendments require that the U.S. Environmental Protection Agency (EPA) study the contribution of nonroad engines to urban air pollution, and regulate them if warranted. "Nonroad," also referred to as "off-road" or "off-highway," includes recreational equipment, farm equipment, construction, farm, and industrial (CFI) equipment, lawn and garden (L&G) equipment, outdoor power equipment, and marine vessels.[10] The CFI and L&G categories are the focus of this paper.

Nonroad equipment emissions for CFI and L&G are estimated using EPA’s NONROAD model.[11,12] Information regarding the specific data used to come up with emission factors associated with this model are not readily available. In particular, the emissions data are from a variety of reports, a number of which 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, the focus here instead was on obtaining as complete of a data set as possible 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 calculated based upon a similar dataset in the NONROAD model

All of the emission factor data that were collected for this work came from technical reports of the California Air Resources Board (CARB), contract reports prepared by Southwest Research Institute (SwRI), a paper published in *J. Air & Waste Manage. Assoc.*, documentation of the NONROAD model, and papers published by the Society of Automotive Engineers (SAE). These data are described in detail by Frey and Bammi (2002a) in the case of L&G data and by Frey and Bammi (2002b) in the case of CFI data.[13,14] The detailed discussion is not repeated here. 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.[15]

For the L&G source category, data for a total of 51 gasoline-fueled engines were included in the database compiled by NCSU.[13] The database includes data collected using the SAE J1088, CARB J1088, and C6M test procedures, and it includes 2-stroke and 4-stroke engines, and handheld and non-handheld engines. However, 2-stroke engines are predominantly used in hand-held applications and 4-stroke engines are predominantly used in non-handheld applications. 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.

For CFI engines, data for a total of 56 engines are included in the final database compiled by NCSU.[14] All but four of these were diesel engines. These data are based upon the 8-, 13-, 21-, and 23-mode tests. 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, most reports had THC and NO_x data in g/hp-hr units only, while only a few had enough additional information provided so that emission factors could be calculated in other units.

RECOMMENDATIONS FOR NONROAD EMISSION SOURCE CATEGORIES

Emission categories are typically defined based upon *a priori* assumptions, rather than based upon data analysis. In the case of nonroad vehicles, it is possible to define so many emission source categories that one is faced with very scarce data from which to estimate emissions for any given category. For example, if *a priori* assumptions are made that there should be many different categories of emission factor for data for different engine size ranges, then it is likely that the already small available data will be subdivided too finely, resulting in very small sample sizes for any one engine size category. To avoid this problem, an empirical approach was used for the development of emission source subcategories within the broader categories of L&G equipment and CFI equipment.[13,14]

The L&G and CFI databases compiled from the available literature were 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, design, and/or size. Two-tailed t – tests for the difference in means at a 5 percent level of significance were done to determine whether there is a statistically significant difference in the mean emission factor estimate for each possible pair of categories. In general, the comparisons were made for six groups of data: for each of the two pollutants (NO_x and THC), three different emission factor units were considered (g/hr, g/hp-hr, and g/gallon).

For L&G equipment, the findings from the comparisons of mean emissions are that there are: (1) significant differences in emissions between 2-stroke engines and 4-stroke engines; (2) significant differences in emissions between handheld and non-handheld engines; (3) no significant differences in emissions between 4-stroke OHV engines and 4-stroke LHV engines; and (4) mixed indications regarding differences in emissions with respect to engine size. In this work, data were divided into two size ranges for 4-stroke engines but not for 2-stroke engines. 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.

For CFI engines, the results of the statistical analysis supported grouping emission factors on the basis of fuel (diesel vs. gasoline) and engine technology (i.e. 2-stroke vs. 4-stroke) in the case of diesel engines. There was no or little evidence to support grouping these data with respect to engine age, engine size, or type of aspiration in the case of diesel engines.

It should be noted that for both L&G and CFI engines, there were insufficient data to evaluate factors such as the influence of deterioration rate, maintenance practices, or ambient conditions with respect to emissions.

METHOD FOR QUANTIFICATION OF VARIABILITY AND UNCERTAINTY IN EMISSION FACTORS

Inter-engine variability in the emission factor data was quantified using parametric probability distribution models. Uncertainty in the mean emission factors, and regarding the fitted distributions, was quantified using bootstrap simulation.

Estimating Inter-Engine Variability

The inter-engine 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 extrapolation to the tails of the distribution beyond the range of observed data. 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. 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 the results of bootstrap simulations. The sensitivity of the uncertainty estimates to different parametric distributions for variability was evaluated. Statistical goodness-of-fit tests typically lack statistical power in situations with small sample sizes; however, statistical goodness-of-fit tests such as the Kolmogorov-Smirnov and Anderson Darling tests were used when fitting distributions to data in the case of the CFI data. However, the insight obtained from the statistical goodness-of-fit tests was typically the same as that obtained based upon graphically comparing the CDF of the fitted parametric distribution to the data, and by visualizing confidence intervals for the CDF estimated using bootstrap simulation. The parametric distributions considered were lognormal, gamma and Weibull. Maximum likelihood estimation was used to estimate the parameters of the fitted distributions. This overall method is similar to that reported by Cullen and Frey.[16]

The basis for selecting parametric distributions includes a combination of theoretical and empirical considerations. Emissions data must be nonnegative and typically are positively skewed. Thus, a symmetric distribution such as the normal distribution will typically be a poor or inappropriate choice to represent variability in emissions. Furthermore, for data sets with substantial variability, a normal distribution fitted to the data will predict the possibility of negative emission values, which is physically impossible. In contrast, the lognormal distribution is non-negative and positively skewed and, therefore, is often useful for representing physical quantities.[17] The lognormal distribution arises as a result of multiplicative or dilution processes and, therefore, is often an appropriate candidate for concentration or emission rate data.[18] The lognormal, gamma, and Weibull distributions are very similar for moderately skewed data sets and therefore can often provide approximately similar fits to data. However, as the skewness of the data increase, the gamma and Weibull distributions are less tail heavy than the lognormal.[16,19] Both the gamma and lognormal distributions have been used by others to describe physical quantities, such as rainfall and pollutant concentrations.[20,21]

Estimating Uncertainty in Average Emission Factors

Bootstrap simulation was used to estimate uncertainty in the average emission factor. Bootstrap simulation is a numerical technique originally developed for the purpose of estimating confidence intervals for statistics.[22] 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.[16,23-26]

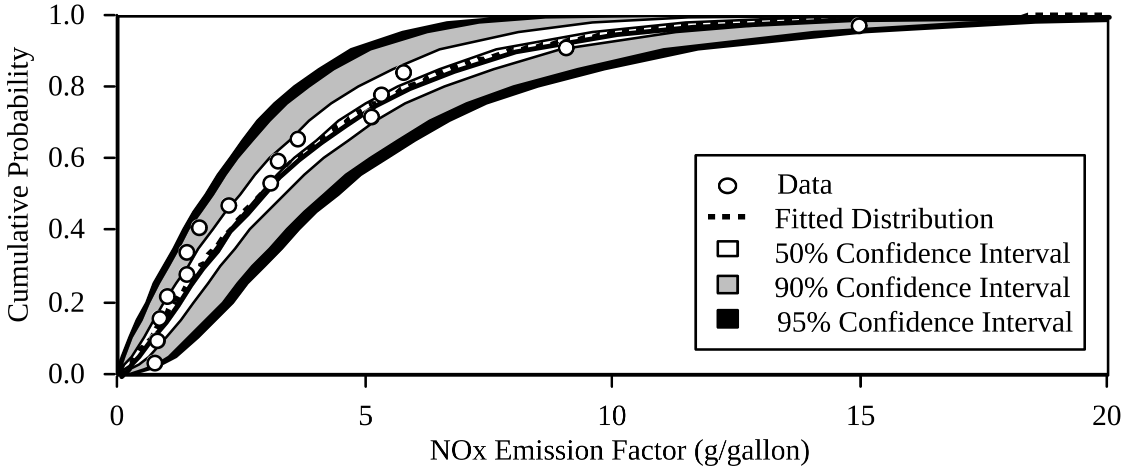
Bootstrap simulation uses a conceptually straightforward approach. First, an estimated population distribution is developed, such as by fitting a parametric distribution to a data set. A random sample of the same size as the original data set is simulated, with replacement, from the assumed population distribution and is referred to as a “bootstrap sample.” Any statistic of interest, such as the mean, standard deviation, distribution parameters, distribution percentiles, or others, is calculated from the bootstrap sample and is referred to as a "bootstrap replication" of the statistic. The process is then repeated many (e.g. 500) 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 2.5th and 97.5th percentiles of the sampling distribution enclose a 95% confidence interval. Confidence intervals were constructed for the mean emission factor estimates and for the fitted cumulative distribution functions. Bootstrap simulations can be repeated a number of times to evaluate numerical stability by comparing results among the multiple bootstrap simulations.

RESULTS OF QUANTITATIVE ANALYSIS OF VARIABILITY AND UNCERTAINTY IN L&G ENGINE EMISSION FACTORS

For each L&G emission factor data set, three candidate probability distribution models for inter-engine variability were fit to the data. These candidates include the lognormal, gamma, and Weibull distributions. In some cases, there is clearly one of the three candidates that provides a better fit to the data than the other two. For example, for less than 8 hp 4-stroke engine NO_x emissions in g/gallon, the Weibull distribution agrees with the data moreso than the other two candidates. The gamma distribution is slightly more tail heavy than the Weibull in this case. The lognormal distribution is noticeably more tail heavy. The gamma distribution agrees more with the 2-stroke NO_x emissions in g/gallon than do the other two candidates, especially in the central portion of the distribution. In some cases, all three distributions provide similar fits. In such cases, there is typically a smaller relative amount of variation in the data than for the other three examples. Thus, as previously noted, there is not much difference between the three distributions for data sets with little or moderate skewness. There is not a single parametric distribution that provides a good fit for all cases. For example, the lognormal distribution does not provide the best fit for all of the emission source categories or emission factor units.

Bootstrap simulation was used to characterize uncertainty in the parameters of the preferred fitted distributions for each emission factor data set. As an example, in Figure 1, probabilistic analysis results are shown for the 2-stroke engine NO_x emission data set in units of g/gallon. The inter-engine variability in emissions is more than an order-of-magnitude, from approximately 1 g/gallon to approximately 15 g/gallon, with most of the data having values of less than 10 g/gallon. A fitted gamma

Figure 1. Fitted Gamma Distribution and Bootstrap Simulation Results for 2-Stroke Lawn and Garden Engine NO_x Emission Data in Units of g/gallon (n=16).



distribution is shown in comparison to the data. The gamma 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 CDF of the gamma distribution are shown. More than half of the data (10 out of 16 data points) are enclosed by the 50 percent confidence interval, and all of the data are enclosed by the 90 percent confidence interval. This comparison suggests that the gamma distribution is an adequate fit to the data set and, therefore, is a reasonable representation of inter-engine variability in emissions.

The results for uncertainty in the mean emission factors for the L&G source categories are summarized in Table 1. Of the 24 cases shown in Table 1, all but one have uncertainty ranges of greater than approximately plus or minus 20 percent, and fourteen have uncertainty ranges of greater than approximately plus or minus 30 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. For example, although the uncertainty in the mean g/gallon emission factor for THC for 2-stroke engine is only plus or minus 12 percent, the range of uncertainty is minus 30 to plus 41 percent if g/hp-hr units are used. There is not conclusive indication that one particular emission factor unit always has less uncertainty than another.

Knowledge of the range of uncertainty enables rigorous comparison among different groups of engines. For example, it is clear that 2-stroke engines have lower average NO_x emissions than the group of all 4-stroke engines because the confidence intervals for the mean emissions for each of the three emission factor units do not overlap. Similarly, it is clear that the group of 2-stroke engines have much higher average THC emissions than the group of all four-stroke engines. It is clear that the smaller 4-stroke engines have significantly higher average THC emissions than the larger 4-stroke engines. The larger 4-stroke engines have significantly higher average NO_x emissions on a g/gal basis, but there is substantial overlap between the confidence intervals of the mean for g/hp-hr units. Thus, it appears to be the case that g/hp-hr units for NO_x vary less for one category to another on a relative basis than do the other two emission factor units.

Table 1. 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	No. of Data	Type of Distrib.	Mean	95% CI on Mean (absolute values)		95% CI on Mean (percentages)	
2-stroke	NO_x	g/gal	16	G	3.7	2.3	5.5	-37	46
		g/hr	18	L	0.93	0.53	1.6	-43	69
		g/hp-hr	16	L	0.83	0.45	1.4	-45	75
	THC	g/gal	16	L	809	720	907	-11	12
		g/hr	18	L	237	182	305	-23	29
		g/hp-hr	16	L	224	156	316	-30	41
4-stroke	NOx	g/gal	19	W	12	8.9	16.2	-28	31
		g/hr	19	L	5.6	2.7	11	-52	100
		g/hp-hr	27	L	2.0	1.5	2.8	-27	35
	THC	g/gal	19	L	113	83	154	-27	35
		g/hr	19	L	34	25	47	-27	36
		g/hp-hr	22	L	22	14	32	-33	46
	NO_x < 8 hp	g/gal	13	W	9.1	6.7	12	-27	29
		g/hr	13	W	1.8	1.3	2.4	-29	32
		g/hp-hr	21	W	1.8	1.3	2.2	-24	26
	NO_x ≥ 8 hp	g/gal	6	L	20	11	36	-47	77
		g/hr	6	W	14	7.7	21	-44	53
		g/hp-hr	6	G	2.7	1.5	4.3	-45	58
	THC < 8 hp	g/gal	13	L	136	101	181	-26	33
		g/hr	13	G	31	21	43	-32	38
		g/hp-hr	16	G	28	19	38	-32	39
	THC ≥ 8 hp	g/gal	6	L	62	40	95	-37	52
		g/hr	6	L	40	27	59	-34	47
		g/hp-hr	6	G	9.0	5.2	14	-42	53

Notes: No. of data = number of data points. Type of Distrib.: G = Gamma; L = Lognormal; W = Weibull. Mean represents the mean of all bootstrap simulations. The 95 % confidence intervals are averages of 15 bootstrap simulations, each based upon 1,000 bootstrap samples. The confidence intervals are reliable to at least two significant figures.

RESULTS OF QUANTITATIVE ANALYSIS OF VARIABILITY AND UNCERTAINTY IN L&G ENGINE EMISSION FACTORS

The methodology for characterizing inter-engine variability and uncertainty in the average emissions for the CFI source categories was similar to that for the L&G source categories. For example, Figure 2 illustrate probabilistic analysis results for 4-stroke diesel engines for the THC emission factor, in units of g/kW-hr. The inter-engine variability in THC emissions ranges from 0.1 g/kW-hr to 5 g/kW-hr. More than half of the data points (23 out of 37) are enclosed within the 50% confidence interval and all of the data are enclosed by the 90% confidence interval. Thus, the Weibull distribution shown in the figure appears to adequately fit the dataset.

Results of the analysis of uncertainty in mean emission factors for all of the cases considered are given in Table 2. For example, for the diesel fueled 4-stroke engine NO_x emissions in units of g/kW-hr, the mean emission rate is 11.3 g/kW-hr. This corresponds to the approximately 55^{th} percentile of the fitted distribution. The mean occurs at a cumulative probability above the median (50^{th} percentile) because the data and the distribution are positively skewed. The 95 percent confidence interval for the mean is 10.2 g/kW-hr to 12.6 g/kW-hr, which is a range of minus 10 percent to plus 11 percent of the mean value. The range is slightly skewed because of the positive skewness in the data set.

Of the 10 cases shown in Table 4, eight have uncertainty ranges of greater than approximately plus or minus 20 percent, and one has an uncertainty range of greater than approximately plus or minus

Figure 2. Fitted Weibull distribution and bootstrap simulation results for 4-stroke diesel-fueled construction, farm, and industrial engine total hydrocarbon emission data.

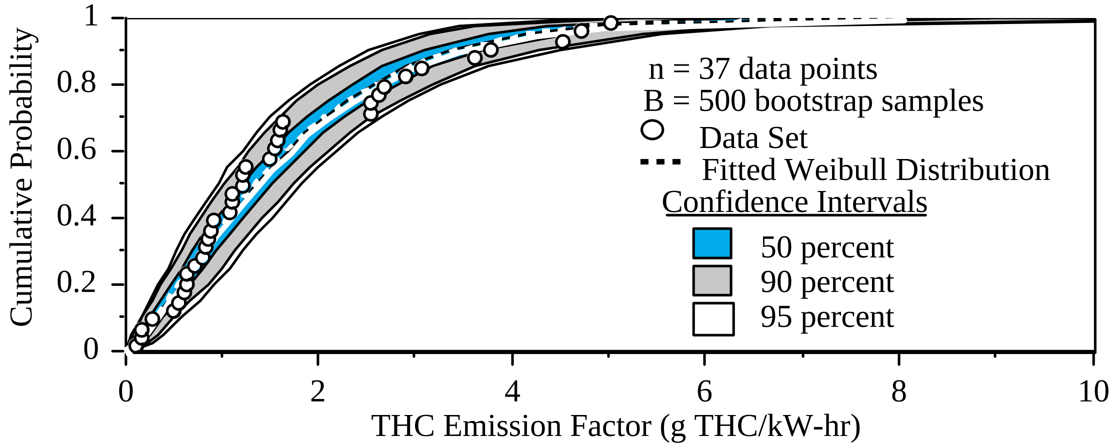


Table 2. Results of the uncertainty analysis of mean NO_x and THC emission rates for gasoline and diesel fueled construction, farm, and industrial engines.

Category	Pollutant	Units	# of Data	Mean	95% CI on Mean	Relative Uncertainty
Gasoline	NO_x	g/kW-hr	4	6.14	4.17 – 8.48	-32% to +38%
	THC	g/kW-hr	4	14.1	11.0 – 16.5	-22% to +17%
Diesel 2-Stroke	NO_x	g/kW-hr	4	22.5	17.7 – 27.1	-21% to +20%
	THC	g/kW-hr	4	2.00	1.05 – 2.97	-48% to +49%
Diesel 4-Stroke	NO_x	g/l	15	38.7	32.8 – 46.0	-15% to +19%
		g/hr	20	1670	1220 – 2140	-27% to +28%
		g/kW-hr	37	11.3	10.2 – 12.6	-10% to +11%
	THC	g/l	15	4.34	2.99 – 5.88	-31% to +34%
		g/hr	20	133	90.3 – 176	-32% to +32%
		g/kW-hr	37	1.68	1.25 – 2.12	-26% to +26%

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 the unit of measurement used for the emission factors, at least in part because differing numbers of data points are available.

CONCLUSIONS

Datasets for L&G and for CFI emission factors were compiled based upon a review of available literature. A key difficulty encountered in this work was to obtain a well-documented and complete database such as that used in EPA's NONROAD model. We recommend that such data should be made widely available through published technical documents and via databases on the World Wide Web.

Measurements of L&G and CFI engines have been made using a variety of test procedures. Care was taken in this work not to group together data that were obtained from dissimilar test methods. For emission inventory purposes, it is important to have data that are representative of real world emissions. 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. For example, on-board emissions measurement methods can be used to measure L&G and CFI equipment emissions during in-use operation.

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 and not based only upon *a priori* assumptions. It is important not to fragment databases by creating too many unnecessary categories, while at the same time it is important to divide the data into categories if there is a good empirical basis for doing so. We recommend that an empirically-based approach to developing emission source categories be employed to avoid creation of unnecessary categories.

Uncertainty in emissions measurements can, potentially, be a significant source of uncertainty. In the judgment of the authors, the methods used to measure engine emissions for NO_x and THC 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 reported information regarding the precision and accuracy of the test methods. A key recommendation is that such information should be provided in emission test reports and in emission factor databases.

This paper has successfully demonstrated, with application to L&G and CFI equipment, a procedure for quantifying variability in emissions and uncertainty in mean emissions factors using parametric probability distributions and bootstrap simulation. 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 was relatively large, with nearly all cases evaluated having uncertainty ranges for the mean in excess of plus or minus 20 percent. The ranges of uncertainty for the mean emission factors are typically positively skewed, which reflects the positive skewness and small sample sizes of the available data.

The large range of quantified uncertainty in emission factors suggests that it is important to quantify uncertainty. As the NRC recently noted, 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. For example, in comparing different engine technologies, it is important to consider the uncertainty in emission factor estimates in order to determine whether one technology has higher or lower emissions than the other with statistical significance. 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/or more data. The probabilistic methodology presented here is part of an overall approach to developing policy, program management, and research planning.

ACKNOWLEDGMENTS

This work was supported by U.S. Environmental Protection Agency STAR Grants Nos. R826766 and R826790. This paper has not been subject to EPA review and no official endorsement by EPA 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 recommendations regarding data sources. We thank Mr. Charles T. Hare, Southwest Research Institute, for providing us with emissions test reports.

KEYWORDS

Nonroad mobile sources, emission factors, uncertainty analysis, bootstrap simulation, Monte Carlo simulation, nitrogen oxides, hydrocarbons, lawn and garden, construction, farm, industrial, engines

REFERENCES

1. National Research Council. *Modeling Mobile-Source Emissions.* National Academy Press: Washington, DC. 2000.

2. US Environmental Protection Agency. *Guiding Principles for Monte Carlo Analysis*, Risk Assessment Forum, Washington, DC, 1997; EPA/630R-97/001.

3. Frey, H.C., N.M. Rouphail, A. Unal, and J.D. Colyar, Emissions Reduction Through Better Traffic Management: An Empirical Evaluation Based Upon On-Road Measurements, FHWY/NC/2002-001, Prepared by North Carolina State University for North Carolina Department of Transportation, December 2001. 323 pp.

4. Frey, H.C., A. Unal, and J. Chen, Recommended Strategy for On-Board Emission Data Analysis and Collection for the New Generation Model, Prepared by North Carolina State University for the Office of Transportation and Air Quality, U.S. Environmental Protection Agency, Ann Arbor, MI. February 2002.

5. US Environmental Protection Agency. *Compilation of Air Pollutant Emission Factors*; AP-42, Fifth Edition; Research Triangle Park, NC, 1995.

6. 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.

7. Dickson, R.J., Hobbs A. D.; "Evaluation of Emission Inventory Uncertainty Estimation Procedures." Paper No. 89-24.8, In *Proceedings of 82nd Annual Meeting*, Air & Waste Management Association: Anaheim, CA, 1989.

8. National Research Council. *Rethinking the Ozone Problem in Urban and Regional Air Pollution.* National Academy Press: Washington, DC, 1991.

9. 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.

10. US Environmental Protection Agency. *Reducing Air Pollution From Nonroad Engines*; at EPA-OTAQ Nonroad Engine Emissions Page. http://www.epa.gov/otaq/regs/nonroad/f00048.htm (accessed January 2001); EPA420-F-00-048.

11. Beardsley, M.; Lindhjem, C. E.; Harvey, C., *Exhaust Emission Factors for Nonroad Engine Modeling - Spark Ignition;* U.S. Environmental Protection Agency, 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); Report No. NR-010b.

12. US Environmental Protection Agency. *Nonroad Engine and Vehicle Emission Study (NEVES)*; NTIS Accession No. PB92-126960, Office of Mobile Sources, Certification Division: Ann Arbor, MI, 1991; EPA 460/3-91-02.

13. Frey, H.C., and S. Bammi, "Quantification Of Variability and Uncertainty In Lawn And Garden Equipment NO_x and Total Hydrocarbon Emission Factors," *Journal of the Air & Waste Management Association*, April 2002 (in press).

14. Frey, H.C., and S. Bammi, "Probabilistic Nonroad Mobile Source Emission Factors," *Journal of Environmental Engineering*, revised and resubmitted March 2002.

15. 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;* Prepared by Department of Civil Engineering, North Carolina State University for North Carolina Department of Transportation, pp 55, 1997; FHWY/NC/97-005.

16. 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.

17. Aitcheson, J.; Brown, J.A.C.. *The Lognormal Distribution*. Cambridge University Press: Cambridge, UK. 1957.

18. Ott, W., "A Physcal Explanation of the Lognormality of Pollutant Concentrations," *J. Air & Waste Manage. Assoc*., 40:1378-1383. 1990.

19. Hahn, G.J.; Shapiro, S.S. *Statistical Models in Engineering*. Wiley Classics Library, John Wiley and Sons: New York. 1967.

20. Eagleson, P.S., "Climate, Soil, and Vegetation," *Water Resources Research*, 14:705-776. 1978.

21. Seinfeld, J.H. Atmospheric Chemistry and Physics of Air Pollution. Wiley-Interscience: New York. 1986.

22. Efron, B.; Tibshirani, R. J. *An Introduction to Bootstrap*; Chapman and Hall: New York 1993.

23. Frey, H.C.; Bharvirkar R.; Zheng J.; 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.

24. 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, Vol. 2(4), 762-797.

25. Rhodes, D.S., and Frey H.C., "Quantification of Variability and Uncertainty in AP-42 Emission Factors Using Bootstrap Simulation," *In Proceedings Emission Inventory: Planning for the Future*, Air and Waste Management Association: Pittsburgh, Pennsylvania, October 1997, pp. 147-161.

26. 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.