Evaluation and Recommendation of a Modal Method for Modeling Vehicle Emissions

H. Christopher Frey, Alper Unal, Jianjun Chen, and Song Li
Department of Civil Engineering
Campus Box 7809
North Carolina State University,
Raleigh, NC 27695-7908
frey@eos.ncsu.edu

ABSTRACT

The Environmental Protection Agency's new Multi-scale Motor Vehicle and Equipment Emissions System (MOVES) is intended to replace the MOBILE6 and NONROAD models in coming years. To establish a methodology for the emission rate estimator portion of MOVES, the following questions were addressed: (1) how should second-by-second data be used to estimate emission rates?; (2) what explanatory variables should be selected to refine the emission rates?; (3) what averaging time is preferred?: (4) what emission factor units should be used?; (5) what weighting approach should be used, when comparing time-, vehicle-, and trip-weighted?; (6) how should variability and uncertainty be characterized?; and (7) how should the conceptual model be validated and what are the results of validation exercises? A modal emissions estimation approach was demonstrated for hot stabilized Tier 1 vehicle emissions based upon binning second-by-second data with respect to 14 vehicle specific power categories. The use of 5 or 10 second averaging times as a basis for estimating modal emission rates did not improve prediction of total cycle or trip emissions. A time-weighted approach, based upon gram per second emission rates, offers the most flexibility for model development. Inter-vehicle variability and uncertainty in mean modal emission rates were characterized and methods for propagating uncertainty in MOVES were recommended. The conceptual approach was evaluated based upon comparison of model predictions to independent data. The predictions were found to be sensitive to a wide range of driving cycles and were accurate in most cases. Opportunities to refine and improve the approach were identified and recommended.

INTRODUCTION

EPA is undertaking an effort to develop a new set of modeling tools for the estimation of emissions produced by on-road and off-road mobile sources. The product of this effort will be the **M**ulti-scale m**O**tor **V**ehicle & equipment **E**mission **S**ystem, referred to as MOVES.[1] The design of MOVES is guided by the following four considerations:

1) the model should encompass all pollutants (e.g., HC, CO, NO_x, particulate matter, air toxics, and greenhouse gases) and all mobile sources at the levels of resolution needed for the diverse applications of the system;
2) the model should be developed according to principles of sound science;
3) the software design of the model should be efficient and flexible; and
4) the model should be implemented in a coordinated, clear, and consistent manner.

A critical element of MOVES is the use of data gathered using on-board emissions measurement devices. To explore this issue, in Fall 2001 EPA issued an on-board emission analysis "shootout" contract in order to solicit several approaches for incorporating on-board emissions into moves. Three shootout contracts were issued to three organizations that worked independently on the same general statement of work.[2-4] Each contractor had the flexibility to choose any approach they preferred. NCSU pursued a modal "binning" approach in which operational bins were defined based on speed, acceleration, and power demand, and refined the estimates within each modal bin using regression analysis. The University of California at Riverside (UCR) pursued a database approach, deriving separate emissions for macroscale, mesoscale and microscale based on a database lookup of individual vehicle and trip results. Environ based their approach on a calculation of vehicle specific power (power per unit mass, or vehicle specific power - VSP), aggregating results over "microtrips" (20 or more seconds, defined by endpoints of stable operation). EPA also developed a conceptual approach based upon binning of data with respect to VSP bins.[5]

The shootout results from NCSU, UCR, Environ, and EPA, revealed several promising approaches for using on-board data in the development of MOVES exhaust emission rates. In particular, the development of modal emission rates using a "binning" approach was successfully demonstrated by NCSU and EPA in the shootout analysis. NCSU tackled the time series nature of the on-board data and illustrated methods for dealing with the data to reduce the influence of the time series. The work by Environ illustrated potential benefits to averaging or smoothing the data. As a result of this work, the proposed design of MOVES is predicated on emission rates defined by vehicle and modal operation "bins," and the development of emission rates for these bins in MOVES is the ultimate purpose of the methodology that will be developed in this project.

The philosophy for MOVES is that it should be as directly data-driven as possible. Some of the key advantages of a data driven methodology include:

- Emission rates can be developed from raw data or based upon summaries of actual data
- Emissions estimates from multiple bins can be weighted to represent any combination of trip and vehicle characteristics.
- Inter-vehicle variability and fleet average uncertainty can be easily estimated based upon appropriate averaging times
- Similar conceptual approaches can be used for different types of vehicles (e.g., on-road gasoline and diesel, nonroad gasoline and diesel) and pollutants (i.e. HC, CO, NO_x, particulate matter, air toxics, and greenhouse gases)
- A modal/binning approach can easily support meso-scale and macro-scale analysis
- Methods have been demonstrated by NCSU for handling cold start emissions as part of the modal/binning approach.
- The modal/binning approaches have been evaluated by validating the approaches in comparison to real-world emission measurements.

A key goal of the binning methodology is to develop modal emission rates in a manner that does not require additional modeling analysis, such as regression modeling, and that eliminates the need for many correction factors common to existing models such as Mobile5 and Mobile6. Ideally, the emission rates estimated for a specific bin should be based directly on the sample of raw data falling into that bin.

On-board data is a promising means for developing tailpipe emissions estimates. However, as noted by EPA and as explained in the NCSU final report from the shootout, in the short-term other sources of data will continue to play an important role in populating or evaluating MOVES.[2,5] Thus, an important step in the development of MOVES is to evaluate the feasibility of techniques for applying the modal binning approach to data from other sources, such as driving cycle dynamometer data and remote sensing device (RSD) data. For example, Frey *et al*. demonstrated an approach for estimating modal emission rates from aggregate data.[2,6]

The key purpose of this project was to evaluate ***methods*** for developing modal emission rates from disparate data sources (e.g., on-board data, laboratory second-by-second data, aggregate driving cycle data, I/M data, and RSD data) for a relatively small "pilot" dataset of light duty vehicles. In the shootout, NCSU demonstrated that similar approaches can be applied to HDDV and to nonroad diesel vehicles; therefore, it was reasonable to focus resources on the example of LDGVs in this project. Furthermore, in previous work, NCSU demonstrated how to develop a bin for cold starts. Therefore, this project focused on hot stabilized tailpipe emissions. This project demonstrated at the proof-of-concept level the methodology for developing modal emission rates in MOVES using a wide variety of data sources, including an evaluation of the applicability of aggregate (bag) data and RSD data. The complete details of this work are in a lengthy project report.[6] This paper will highlight the development of model emission rate bins and will summarize the key recommendations. Because the focus of this work is *methodological*, it was *not* the objective of this work to develop a definitive all-inclusive model for all possible vehicle technologies, model years, and so on.

TECHNICAL APPROACH

The objectives of this project were as follows:

- Develop, demonstrate, and report an analytical *approach* for producing exhaust modal emission rates and emission rate distributions for MOVES from a variety of data sources, possibly including aggregate (bag) data and RSD data.
- Develop, demonstrate, and report a *methodology* for estimation of model uncertainty and variability in emissions estimates
- Validate the developed approach by comparing a conceptual model against an independent dataset
- Develop a recommended step-by-step methodology for generating modal emission rates in MOVES.

The key starting point of the work was the development of an analysis data set that included on-board data, second-by-second laboratory data, IM240 data, aggregate (bag) data, and RSD data. Alternative binning approaches were evaluated, averaging times were compared, emission factor units were compared, and the methods for weighting of data were evaluated. The choice of averaging time and of weighting method influenced the results obtained for the uncertainty analysis method.

The following key questions emerged and were addressed during the project:

1. What dataset should be used for the final version of the conceptual model?
2. Which binning approach should be used?
3. How much detail should be included in the binning approach, in terms of how many explanatory variables and how many strata for each variable?
4. What averaging time is preferred as a basis for model development?
5. What emission factor units should be used?

6. What weighting approach should be used, when comparing time-weighted, vehicle weighted, and trip weighted?
7. How should variability and uncertainty be characterized?
8. How should aggregate bag data be analyzed to derive estimates of modal emission rates?
9. What is the potential role and feasibility of incorporating RSD data into the conceptual modeling approach?
10. How should the conceptual model be validated and what are the results of validation exercises?

The main focus of this paper is with regard to Questions 1 and 2. Although there is not sufficient space in this paper to give details regarding how each of these questions are answered, the results and recommendations for each of these ten questions are summarized.

DEVELOPMENT OF ANALYSIS DATASET

A combined data set for running exhaust emission rates for LDGV was developed based upon data provided by EPA, including:

- Approximately 100,000 seconds of data from 17 on-board vehicles from the "shootout" analysis;[8]
- Approximately 75,000 seconds of data on 25 vehicles tested at EPA's lab for the Mobile6 facility-specific driving cycles and other standard cycles;
- 82,800 seconds on 311 vehicles tested on the IM240 as part of the Colorado IM program;
- Bag-only and second-by-second data on 74 vehicles tested over FTP (i.e., Bag 2 and Bag 3) and US06 for development of UC Riverside's Comprehensive Modal Emission Model; and
- RSD data on 200,966 Tier 1 LDGVs collected as part of a remote sensing study in Missouri.

Several quality assurance and post-processing steps were applied to the dataset and are detailed elsewhere.[6,7] Variables such as acceleration, power demand and Vehicle Specific Power (VSP) were estimated from measured variables such as vehicle speed.[9,10] While it is recognized that the estimate of VSP is a function of vehicle weight and of the specific values of the parameters for each individual vehicle, it was beyond the scope of this study to develop detailed vehicle-specific estimates of VSP. Instead, VSP estimated based upon average or typical values of these coefficients is used here as a means to estimate a single metric based upon vehicle speed, acceleration, and road grade. The usefulness of this metric as an explanatory variable is evaluated via statistical analysis.

The combined database was used to create specific databases for different analyses throughout the project. These databases included the following:

- A "Modeling" or "Calibration" database comprised of data for most of the on-board measurements, most of the EPA dynamometer data, and most of the NCHRP data. This database was also used as "Validation Data Set 1"
- "Validation Data Set 2" was comprised of a small sample of vehicles from the EPA on-board, EPA dynamometer, and NCHRP data that were excluded from the modeling database.
- IM240 data were used separately from the other data

- The NCHRP data were used in the analysis of methods for developing modal emission rates from aggregate bag data
- “Validation Data Set 3” was comprised of data obtained from the California Air Resources Board, and are also referred to as “ARB data.”
- RSD data included approximately 2,000,000 seconds of data. Of this dataset, 200,966 data points were selected randomly for analysis, where each point represents measurement for one vehicle.

The data from on-board, EPA dynamometer and NCHRP dynamometer measurements were combined into the modeling data set, and included:

- 71,699 seconds of data from 13 on-board vehicles from the “shootout” analysis;
- 68,482 seconds of data on 33 vehicles tested at EPA’s lab for the Mobile6 facility-specific driving cycles and other standard cycles; and
- 92,000 seconds of data on 49 vehicles tested over FTP and US06 for development of UC Riverside’s Comprehensive Modal Emission Model.

Therefore, the combined database for modeling had a total of 232,181 seconds of data. The combined database has the following data fields: source for data (e.g., EPA dynamometer); vehicle make; vehicle model; VIN; number of vehicle tested; number of trip tested; speed; acceleration; ambient temperature; ambient humidity; road grade; estimated power, estimated positive power; estimated VSP; estimated positive VSP; CO, CO_2, HC, NO_x emissions; vehicle model year; vehicle engine displacement; number of cylinders; air condition use; and vehicle net weight.

Validation Data Set 2 included the following data:

- 3 vehicles from EPA dynamometer data
- 3 vehicles from EPA On-board data
- 25 vehicles from NCHRP data

The validation dataset included 83,183 seconds of data. The data fields for this dataset were the same as for the Modeling dataset.

The NCHRP dataset included 8 high-emitter vehicles as reported in a User’s Manual prepared by the University of California at Riverside. In preparing Validation Dataset 1 and 2, data were selected randomly from NCHRP data. Six of the high emitter vehicles were included in Validation Dataset 1, and two of them were included in Validation Dataset 2.

Validation Data Set 3 included data for 17 vehicles from 11 different UCC cycles. The validation dataset included nominal speed profiles and total emissions for 15 of the vehicles, and actual speed profiles and second-by-second emissions for two of the vehicles. Detailed information regarding the Validation Datasets is given in the Appendix.

Data for IM240 were utilized for comparative purposes, including comparing average emission rates for the developed modes with respect to those obtained from the calibration data. The IM240 dataset included 311 vehicles tested on the IM240 cycle, for a total of 82,800 seconds of data.

EPA obtained an RSD database from the state of Missouri that contained approximately 2 million records. Of this dataset, 200,966 data points were selected randomly for analysis. Vehicle net weight was not available and engine displacement was only available for part of the dataset. Each data point in the RSD database used for analysis represents a unique vehicle.

DEVELOPMENT OF A MODAL EMISSIONS MODELING APPROACH

The objective of this section is to demonstrate the modal "bin" approach on data for "running" hot-stabilized exhaust emission rates. This section focuses upon the use of one second data in units of mass per time. As part of other work, different averaging times and emission factor units were compared.[6] The two most promising binning approaches identified in the "shootout" were the VSP-based approach evaluated by EPA and the driving mode-based approach evaluated by NCSU. These two approaches were compared in this project; however, the focus here is on the VSP-based approach. A key methodological component of this work was the use of Hierarchical Tree-Based Regression (HTBR), using S-Plus software. This section focuses on answering the second key question of this project: which binning approach should be used? First, the methodology for developing bins based upon statistical methods is presented. Results of analysis of the modeling data set based upon the VSP based approach are presented.

Statistical Method for Developing Binning Criteria

HBTR is a forward step-wise variable selection method, similar to forward stepwise regression. This method is also known as Classification and Regression Trees (CART). Conceptually, HTBR seeks to divide a data set into subsets, each of which is more homogeneous compared to the total data set. At a given level of division, each of the subsets is intended to be different in terms of the mean value. Thus, HTBR is a statistical approach for binning data.[11] The iterative partitioning process is continued at each node until one of the following conditions is met: (1) the node of a tree has met minimum population criteria which is the minimum sample size at which the last split is performed; or (2) minimum deviance criteria at a node have been met.

In developing bins, vehicle-based variables such as vehicle class, mileage, age, engine size, vehicle weight, and technology were utilized. Vehicle operation variables such as vehicle speed, acceleration, and VSP were included in this analysis. Based upon the availability of the data, external parameters such as road grade, air condition usage, ambient temperature, relative humidity were incorporated during HTBR analysis.

In developing bins both "unsupervised" and "supervised" techniques were utilized. In the "unsupervised" technique, data is provided to the HTBR with no prior specification of branches or nodes of the regression tree. In this situation, HBTR is allowed to create whatever bins result from direct application of HBTR. In contrast, for the "supervised" technique, HTBR is forced to start with pre-determined modes. A partially supervised technique can often be a better approach than a purely unsupervised technique. This is because HTBR can be sensitive to artifacts of variability in the data that may not be important from a practical perspective, and HBTR may give unexpected or difficult to interpret results if the unsupervised technique is used. Sometimes HBTR will repeatedly "split" on the same subset of variables (e.g., speed and acceleration) which may indicate the need for a new explanatory variable that is a function of the subset of variables. For example, if HBTR splits repeatedly on speed and acceleration, it may be better to remove speed and acceleration as criteria for creating bins and instead offer some variable that is a combination of both speed and acceleration, such as VSP or power demand.

There is a trade-off between the number of bins and the usefulness of the empirical model based upon the bins. While it is possible to obtain additional explanatory power by increasing the number of bins, there are diminishing returns associated with creation of an increasing number of bins. Furthermore, the HBTR determines bins based upon whether there are differences in the average emissions among the possible bins. It does not determine bins based upon what portion of trip or total emissions are explained by each bin. Therefore, it is possible to obtain a potentially large number of bins that do not help explain a significant portion of total trip or aggregate emissions.

The key explanatory variables were identified by using HBTR and identifying which variables were selected for the first several splits of the data. These variables were selected as the basis for defining modes. The percent contribution of each candidate mode to total emissions was estimated. Based upon these results, the definitions of the modes were revised so that no single mode contributed disproportionately to the total emissions represented in the database.

Application of Statistical Methods to Selecting a Basis for Defining Modes

HBTR was applied to the modeling dataset in order to determine whether VSP would be selected by HTBR as the most important explanatory variable. Vehicle operating parameters as well as vehicle technology parameters were used as possible explanatory variables. These parameters are: speed; acceleration; VSP; temperature; engine displacement; number of cylinders; a/c usage; temperature, odometer reading; model year; and net weight. Of all these parameters VSP was selected as the first split by HTBR. Under the second branch of the tree, a second split was made based upon vehicle net weight. However, the reduction in deviance based upon further stratification by net weight is less than the reduction in deviance from the first split based upon VSP. At the lowest portion of the tree, a second split based upon VSP was observed for the smaller net weight category of data. When a variable occurs repeatedly in the tree, such as VSP does in this case, that is evidence that the variable plays an important role. In this case, VSP alone helps explain a substantial portion of deviance in the data. When the data are further stratified, VSP explains additional deviance for vehicles with a net weight less than 4,400 pounds. This result illustrates that VSP is the most important variable and therefore could be selected as the first criteria for developing bin definitions. Qualitatively similar results were obtained for other pollutants.

Because VSP was consistently identified as the most important explanatory variable for each pollutant, modal bins were developed using VSP. A "supervised" approach was used to develop the final definitions of the bins, taking into account two key considerations: (1) each mode should have a statistically significantly different average emission rate than that of any other mode; and (2) no single mode should dominate the estimate of total emissions for a typical trip as represented by the database. Therefore, to guide the selection of modal definitions, it was decided that no mode should explain more than approximately 10 percent of total emissions. The same modes were defined for all the pollutants. Table 1 provides the mode definitions. Figure 1 shows average modal rates for these bins for all four pollutants. The average modal rates for the

Table 3-1. Definitions for VSP Modes

VSP Mode	Definition
1	VSP<-2
2	-2<=VSP<0
3	0<=VSP<1
4	1<=VSP<4
5	4<=VSP<7
6	7<=VSP<10
7	10<=VSP<13
8	13<=VSP<16
9	16<=VSP<19
10	19<=VSP<23
11	23<=VSP<28
12	28<=VSP<33
13	33<=VSP<39
14	39<=VSP

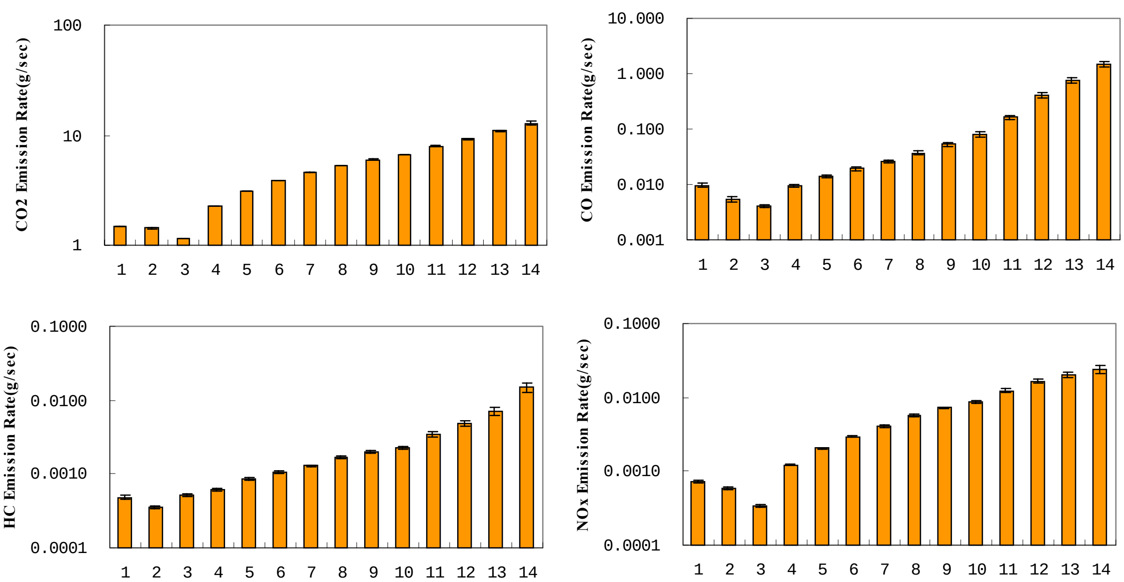

Figure 1. Average Modal Emission Rates for 14 Vehicle Specific Power Bins for CO_2, CO, Hydrocarbons (HC), and NO_x Based Upon the Modeling Data Set.

first two modes, Modes 1 and 2, are higher than average rate for Mode 3 except for HC. There is an increasing trend in emissions with increase in VSP bins for Modes 4 through 14 for all of the pollutants. For CO, the range in average modal emissions is more than two orders-of-magnitude, when comparing Mode 3 and Mode 14, whereas the range is approximately one to two orders-of-magnitude for NO_x, HC, and CO_2.

Because each pollutant has a different sensitivity to the modal definitions, there are some cases in which a mode may contribute approximately 10 percent to the total emissions of one pollutant but a far lower percentage of total emissions for another pollutant. For example, for the high

VSP bins, such as Modes 12, 13, and 14, approximately 10 percent of the total CO emissions in the calibration data set are accounted for, for a total of over 30 percent of the total CO emissions. These four modes account for less than three percent of total travel time in the database. Furthermore, these modes account for only approximately 15 percent or less of total NO_x, HC, and CO_2 emissions. The implication is that high VSP has a more substantial impact upon CO emissions than for the other pollutants. This seems plausible, in that high VSP is likely to be associated with an increased frequency and duration of command enrichment, which tends to have more effect on CO emissions than, for example, NO_x emissions. Of course, the proportion of emissions in each mode is conditional on the database used to estimate the modal emission rates.

SUMMARY OF MAIN FINDINGS

The answers to the ten questions posed in the technical approach are briefly summarized here. More detail regarding how answers were developed for these questions is given in a separate technical report.[6]

Question 1: What dataset should be used for the final version of the conceptual model? The data set used for the conceptual model was comprised of EPA dynamometer data, EPA on-board data, and NCHRP dynamometer data for Tier 1 vehicles only. These data comprised the modeling database. The modeling database was compared to several other databases, including an IM240 database and an RSD database. Since the objective of this work was to develop a methodology, and not to develop a final model, it was not necessary to consider all possible vehicle types. However, clearly, it will be necessary to apply the methodology developed here to vehicles other than Tier 1 in the future.

Question 2: Which binning approach should be used? The binning approach selected was a 14 mode VSP-based approach. However, it was shown in separate analyses that an approach based upon driving modes of idle, acceleration, cruise, and deceleration produced comparable predictions for total emissions. Thus, the 14 mode VSP-based approach is not unique in its capability to predict emissions, but it is expected to facilitate design of a modeling system perhaps moreso than the other approach. The selection of 14 modes is based upon a specific selection criteria. A different number of modes would be implied by other criteria.

Question 3: How much detail should be included in the binning approach, in terms of how many explanatory variables and how many strata for each variable? There is a trade-off between improving the explanatory power of a model and having a model that becomes complicated to code or use. Odometer reading and engine displacement were identified as key explanatory variables. Engine displacement is highly correlated with vehicle net weight and with the number of cylinders of the engine. Therefore, it is not necessary to include net vehicle weight or number of cylinders if engine displacement is selected as an explanatory variable. The objective of this work was to develop a *methodology* for defining bins, but that it was not an objective of this work to apply the methodology to all possible vehicle categories. The latter is a recommendation for future work. Clearly, model year and/or other indicators of different vehicle technologies will be important when the method is extended to other than just Tier 1 vehicles. Ambient parameters such as humidity were accounted for in correcting NO_x emissions. Ambient temperature was not found to be a significant explanatory variable. This result could be different for other vehicle technologies or for perhaps wider ranges of variation in ambient temperature than were observed in the modeling data base. There may be an opportunity to improve the explanatory power of the 14 mode VSP-based approach by including either speed or acceleration as a criteria for further disaggregating the bins. In preliminary analyses it was found that there are differences in average emissions for a given VSP bin when the data in the bin are further stratified with respect to speed and/or acceleration. However, such differences did not substantially improve comparisons of predicted total trip emissions with independent data.

An approach based upon "56 bins" for which the 14 VSP modes were stratified into two odometer reading categories and two engine displacement categories performed reasonably well when predictions were compared to observations for independent data sets. The validation case studies thus emphasize that the modal emissions approach is feasible. A key benefit of the conceptual modeling approach is that it works for all four pollutants considered, and it is not necessary to develop a separate approach for each pollutant.

Question 4: What averaging time is preferred as a basis for model development? Three averaging times were compared with respect to ability to make predictions of trip emissions. No substantial difference was found. Thus, for simplicity, the one second averaging time was recommended for model development and was employed in this work. However, although the issue of averaging time may not have a significant effect on prediction of average emissions, there is a significant effect on the prediction of uncertainty in average emissions. The range of uncertainty in the average modal emission rates is a function of averaging times, and the uncertainty estimates should be adjusted appropriately when making predictions of uncertainty.

Question 5: What emission factor units should be used? With regard to emission factor units, there was no clear overall advantage for emission ratios of CO/CO_2, HC/CO_2, and NO_x/CO_2 versus mass per time emission factors for CO, HC, and NO_x. Although it is the case that there is less variability in the averages among many of the modes for CO and HC for emission ratios when compared to mass per time emission rates, for NO_x there is substantial variability across all modes regardless of the units used. For software design purposes, it is simpler to use the same approach for all pollutants. Thus, an emission ratio approach would require a similar number of modes as the mass per time approach. Additionally, it is necessary to estimate mass per time emissions of CO_2, or to estimate mass per time fuel consumption, in order to convert emission ratios for CO, HC, and NO_x to mass emission rates as would be required for an emission inventory model. Although an emission ratio approach offers some benefits of simplicity when applied to an areawide macroscale emission inventory based upon information such as fuel sales, an emission ratio approach nonetheless would require modal estimates of CO_2 emissions or fuel use when applied to mesoscale emission inventories. Thus, for consistency in the modeling approach, taking into account the objectives of the desired model, the preferred strategy was to use mass per time emission rates for all pollutants and to apply the same modal emissions approach for all pollutants.

Question 6: What weighting approach should be used, when comparing time-weighted, vehicle weighted, and trip weighted? Three weighting approaches were compared, including time, trip, and vehicle weighted approaches. It is clear that the average emission estimates will differ depending on which approach is used, because each approach gives a different amount of weight to different subgroups of the data. For example, the time weighted approach gives equal weight to each data point. The trip weighted approach gives each trip (or driving cycle test) equal weight, even though trip lengths may differ and even though some vehicles may be represented by many trips and others may be represented by only one. The vehicle weighted approach gives each vehicle equal weight regardless of the total testing time or number of trips (or tests). When comparing time, trip, and vehicle weighted approaches, the standard deviation of the variability in emissions decreases in the same order because each successive approach involves more averaging. However, the averaging time is not standardized for the trip and vehicle weighted approaches. Because averaging time is important to accurate estimation of uncertainty, preference was given in this work to the time weighted approach. Note also that if RSD data were to be incorporated and if a vehicle weighted approach were used, then a single RSD measurement of less than one second would be of comparable importance to hundreds or thousands of seconds of measurements from dynamometer or on-board measurements. Clearly, the informational value of a 0.6 second RSD measurement is not the same as from measurements over an entire trip or cycle.

Question 7: How should variability and uncertainty be characterized? The recommendations regarding these issues are given in more detail elsewhere.[6] In brief, the feasibility of representing variability in modal emission rates with parametric distributions was demonstrated. In some cases, single component parametric distributions cannot provide a good fit, but in such cases a two component mixture of lognormal distributions provided an excellent fit. The Method of Matching Moments is recommended as a preferred parameter estimation method if the objective is to have the mean and standard deviation of the fitted distributions match those of the data. For mixture distributions, MoMM is not considered a feasible parameter estimation method and Maximum Likelihood Estimation is recommended. However, the differences in results between MoMM and MLE become smaller as the goodness-of-fit improves. Thus, a well fitting mixture distribution will typically have a mean and standard deviation similar to the data.

The analysis of uncertainty need not be conditioned upon the assumptions made regarding the characterization of variability based upon parametric distributions. For example, uncertainty in the mean can be estimated directly based upon the data using analytical or numerical methods. It is recommended that the sample size and the relative standard error of the mean of each bin be quantified. If the sample size is less than 40 and/or if the relative standard error of the mean is greater than 0.2, then bootstrap simulation is recommended as a technique for quantifying the sampling distribution of the mean. In all other cases, a normality assumption will typically be more than adequate. Parametric distributions can be fit to sampling distributions obtained from bootstrap simulation. Thus, for all modes, it is possible to use parametric distributions to represent uncertainty in the mean, which will facilitate software design and model applications.

Both numerical and analytical methods for propagating uncertainty through a model were explored. Numerical methods such as Monte Carlo simulation or Latin Hypercube Sampling offer the advantage of increased flexibility to accommodate many kinds of distributions and models, including situations in which uncertainty is quantified not only for modal emission rates but also for vehicle activity. In contrast, the analytical approach offers the advantage of less computational burden but is less flexible. An exact solution can be obtained for linear combinations of normal distributions, such as when uncertainty in only modal emission rates is quantified and when all such uncertainties are assumed to be normally distributed.

The range of uncertainty in total emissions estimates was large enough in many cases to justify the importance of performing an uncertainty analysis. For example, for HC and CO emissions the range of uncertainty was as large as plus or minus 30 percent for selected vehicle groups and for four different driving cycles.

Question 8: How should aggregate bag data be analyzed to derive estimates of modal emission rates? With respect to the issue of how to estimate modal emission rates from aggregate dynamometer data (for which no second-by-second data are available), the results were mixed. It is possible to develop good modal emission estimates especially for CO_2 as long as there is a sufficient sample size and as long as sufficient constraints are specified in the least squares optimization approach. However, the range of uncertainty in the predicted modal emission rates can be much larger than the uncertainty in modal emission rates obtained from second by second data. Thus, it is important to develop good estimates of the constraints. However, when applied to vehicle groups for which there are no or few comparable second-by-second data, such as for older carbureted vehicles, it may be difficult to develop good estimates of the constraints. An alternative approach is to arbitrarily specify more stringent constraints, such as defining ratios to be multiples of each other, in which case the estimation problem becomes simpler but the answers obtained will be highly conditional upon such constraints.

Question 9: What is the potential role and feasibility of incorporating RSD data into the conceptual modeling approach? A critical issue for any model is to have a representative data set for calibration. A representative data set should have proportional representation of vehicle emission rates and activity patterns similar to that in the real world. The modeling database was compared to IM240 and remote sensing data. The modeling database produces lower emissions estimates for some modes and comparable emissions estimates for others when compared to these other data sources. A possible reason for the differences could be because of a different representation of high versus normal emitting vehicles. However, the activity patterns of the modeling database are generally different than those of the IM240 and RSD data. Thus, a key question is not only whether the modeling database sufficiently represents high emitting vehicles, but whether the IM240 and RSD data adequately represent real world activity patterns from which it is useful to make inferences regarding emissions. The evidence to support an answer to this question is inconclusive given the different activity patterns for the IM240 and RSD databases compared to that of the modeling database, as well as the possibility of other potential confounding factors, such as fuel effects. From a methodological perspective, the main implication of these comparisons in terms of future model development is to make sure that the modeling database for future work is more comprehensive in terms of sample size and coverage of vehicles considered to be both normal and high emitters.

IM240 data do not contain sufficient variability in driving conditions to fully represent real world driving. For example, IM240 data are limited with respect to the range of VSP compared to real-world on-road trips for which on-board data were available. RSD data are site-specific, and often sites are selected that have only one lane of travel and include a positive road grade. Therefore, RSD measurements are lacking for many locations of real world interest, such as at or near intersections or across multiple lanes of heavy traffic. Additional assumptions have to be made in order to convert these estimates into mass per distance or mass per time emission factors. The apparently large sample sizes for RSD data are for very short averaging times, and thus the informational content of such large sample sizes is not directly comparable to that of time series measurements on a smaller number of vehicles obtained from other methods that include more variation in activity.

Question 10: How should the conceptual model be validated and what are the results of validation exercises? Three approaches were taking toward validation of the conceptual modeling approach. The first was to perform a consistency check, which demonstrated that the modal emission approach can be applied to a dataset to disaggregate emissions into modes, and that it is possible to reaggregate the model emissions and reproduce the total trip emissions.

The second validation case study was to compare model predictions to observed values for a set of vehicles similar to but not identical to those used in the modeling data base. The comparison demonstrated that differences in vehicle mix between the modeling database and the validation database can lead to differences when comparing predicted and observed emissions. However, for cases in which the model and the observed values agreed well for CO_2 emissions, they also tended to agree well for emissions of the other three pollutants. In the future, it is worthwhile to perform similar validation studies by withholding data from the modeling database for some of the trips made by a subset of vehicles, rather than to withhold from the modeling database all data for a particular set of vehicles. Such an approach would improve the likelihood that the vehicles in the validation data set are similar to those in the modeling data set.

The third validation case study involved prediction of emissions for an independent set of vehicles based upon data provided by CARB. The comparison of predicted and observed emissions was generally excellent for CO_2, NO_x, and CO for eight different driving cycles. The model overpredicted for HC in all cases; however, it is possible that CARB may have reported only nonmethane hydrocarbons instead of total hydrocarbons or that there was a fuel effect. A potential distinction between Tier 1 and TLEV vehicles in the CARB database was explored. However, no significant difference in emissions was found for vehicles that might be TLEVs versus those that were Tier 1; therefore, it was not useful to report results separately for these two possible categories.

CONCLUSIONS

In conclusion, this work has demonstrated the feasibility of an empirically-based method for modal emissions model. The methods demonstrated in this work should be used or adapted for use in the development of MOVES and other emission estimation systems.

ACKNOWLEDGMENTS

This work was supported by the Office of Transportation and Air Quality of the U.S. Environmental Protection Agency. This paper has not been subject to EPA review and therefore no official endorsement should be inferred. The authors are grateful to John Koupal, David Brzezinski, James Warila, Connie Hart, Chad Bailey, Mitch Cumberworth, and Craig Harvey of EPA/OTAQ and to Edward Nam of Ford Motor Company for their input during the course of the work.

REFERENCES

1. EPA (2002), Draft Design and Implementation Plan for EPA's Multi-Scale Motor Vehicle and Equipment Emission System (MOVES)," EPA420-P-02-006, U.S. Environmental Protection Agency, Ann Arbor, MI, October 2002.
2. NCSU (2002a) "Recommended Strategy for Onboard Data Analysis and Collection for the New Generation Model" Prepared by North Carolina State University for U.S. Environmental Protection Agency, Ann Arbor, MI, February 4, 2002.
3. UC-Riverside (2002), "Mobile Source Emissions New Generation Model: Using a Hybrid Database Prediction Technique" Prepared by University of California at Riverside for U.S. Environmental Protection Agency, Ann Arbor, MI, February 6, 2002.
4. ENVIRON (2002) "Onboard Emission Data Analysis and Collection for the New Generation Model" Prepared by ENVIRON Internation Corporation for U.S. Environmental Protection Agency, Ann Arbor, MI, February 22, 2002.
5. EPA (2002b) "EPA's Onboard Analysis Shootout: Overview and Results," EPA420-R-02-026, U.S. Environmental Protection Agency, Ann Arbor, MI, October 2002.
6. NCSU (2002b), "Methodology For Developing Modal Emission Rates For EPA's Multi-Scale Motor Vehicle & Equipment Emission System", EPA420-R-02-027, " Prepared by North Carolina State University for U.S. Environmental Protection Agency, Ann Arbor, MI, August 31, 2002.
7. 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
8. Sensors (2002), "On-Road Emissions Testing of 18 Tier 1 Passenger Cars and 17 Diesel Powered Public Transportation Buses," EPA420-R-02-030, Prepared by Sensors, Inc. for U.S. Environmental Protection Agency, Ann Arbor, MI, October 22, 2002.
9. Bachman, W. H. (1999), *A GIS-Based Modal Model of Automobile Exhaust Emissions Final Report*; Prepared by Georgia Institute of Technology for U.S. Environmental Protection Agency; Atlanta, Georgia.
10. Jimenez, Ph.D. Dissertation, Massachusetts Institute of Technology, Cambridge, MA
11. Washington, S., J. Wolf, and R. Guensler (1997), "A Binary Recursive Partitioning Method for Modeling Hot-Stabilized Emissions from Motor Vehicles," Prepared by School of Civil and Environmental Engineering, Georgia Institute of Technology for the 76th Annual Meeting of the Transportation Research Board, Atlanta, Georgia.