PERFORMANCE EVALUATION OF MEASURE EMISSION RATES: A COMPARISON TO MOBILE5A

Ignatius Fomunung, Simon Washington, Randall Guensler, and William Bachman

ABSTRACT

For the past five years, the Georgia Tech Research Partnership has been developing a research-grade motor vehicle emissions model within a geographic information system (GIS) framework. The model has been developed under a cooperative agreement with the US Environmental Protection Agency (USEPA) and the Federal Highway Administration (FHWA), and has involved the direct funding and in-kind contributions from a variety of public and private research partners. MEASURE (Mobile Emissions Assessment System for Urban and Regional Evaluation) is a comprehensive modeling package that integrates predictive modules for vehicle fleet composition, vehicle activity, onroad modal operation, and vehicle emission rates. The GIS framework of the model allows the linkage of typical 4-step travel demand model outputs, simulation model outputs, or monitored ATMS traffic volume estimates. The MEASURE model also contains several new modal emissions modeling approaches designed to improve predictions of CO, HC, and NOx for the onroad vehicle fleet. Hot stabilized emission rate algorithms within the MEASURE predict emission rates for motor vehicle technology groups as a function of the conditions under which the vehicles are operating, specifically speed and acceleration profiles. Vehicle activity variables modeled in MEASURE include average speed, acceleration rates, deceleration rates, idle time, and surrogates for power demand imposed on the engine. Vehicle and emission control technology variables include fuel metering system, catalytic converter type, supplemental air injection, and transmission speed.

Although previous MEASURE-related papers have detailed model estimation methods, to date the authors have not reported the results of model validation. Previous publications have also never offered a comparison of the MEASURE model emission rate algorithms to those employed in the USEPA's MOBILE5a emission rate model. The validation efforts reported in this paper test of the ability of the MEASURE statistical model to adequately represent emissions cause-effect relationships and compare the MEASURE validation results with those of MOBILE5a. The statistical comparisons show that the MEASURE model provides significant improvements in both average emissions estimates and explanatory power over MOBILE5a for all three pollutants across almost every operating cycle tested. The most significant improvements arise in the CO and HC estimates, but even NOx emissions, which are a function of average speed, improve under the MEASURE modeling regime. In addition, the MEASURE model appears to be less biased (the most critical model performance measure for point-estimate forecasts) than MOBILE5a.

INTRODUCTION

Uncertainties in current emission rate models arise in part because the models rely primarily on average speed as the dominant continuous independent variable in the regression. However, many factors (both continuous and discrete), in addition to average speed, affect the net load demanded of an engine, which in turn affects vehicle emissions. These factors include roadway grade, rolling resistance, aerodynamic drag, engine speed, engine friction losses, transmission losses, mass of vehicle, power consumption of accessories, etc. Numerous references identify these factors as influential in the formulation of various pollutants. However, these factors are largely omitted in current emission prediction algorithms (Guensler, 1993).

Cicero-Fernandez, et al. (1996, 1997a), demonstrated that emissions from an individual vehicle may increase by a factor of two when driven on an uphill grade, yet current inventory models do not account for grade. In addition, real world driving conditions, in terms of speed/acceleration distributions and/or traces, are not well represented in the current models. The Federal Test Procedure (FTP), used appropriately to develop baseline emission rates, does not capture the extremes of emission producing activities associated with aggressive driving. Jimenez-Palacio (1999), using a new definition of specific power, calculated the maximum specific power of the FTP to be approximately 22kW/metric ton. More telling, the research indicates that the onset of commanded enrichment for many vehicles occurs at this maximum. Commanded enrichment is responsible for elevated or “super” emissions, which can be one to several orders of magnitude higher than emissions obtained under stoichiometric engine operation. As a result, a large proportion of commanded enrichment is not likely to appear under the FTP. The supplemental Federal Test Procedure (which will put laboratory vehicles under a significantly increased load) will help alleviate the disconnect between laboratory and onroad operations. Nevertheless, it is critical to ensure that onroad emissions modeling can address the higher emission rates noted to occur in real-world conditions.

Driver behavior may also be an important source of uncertainty and variability in motor vehicle emissions (Bishop, et al., 1996). A study of repeated measurements on the IM240 driving cycle indicated that driver behavior may be responsible for potentially order-of-magnitude differences in emissions for clean low-emitting vehicles (Webster and Shih, 1996). Despite this recognition, few advances have been made in quantifying the effect of driving behavior on emissions, save for Shih, et al., (1997) who used throttle position distributions to represent driver behavior, albeit with mixed results. Their research provides evidence that throttle position distributions might be used to reflect differences in driving behavior, but such models still need refinement, and forecasting throttle position distributions (and interactions with specific driver types, facility type, and trip purpose) may prove too difficult to result in a useful model.

Emerging Models

Efforts at improving motor vehicle emissions have captivated researchers for quite some time. Cadle, et al., (1998) recently summarized advances in real-world motor vehicle emissions modeling. The USEPA is currently revising the MOBILE5a emission rate model. MOBILE6 promises to provide significant improvements in terms of representing modal impacts on emission rates because supplemental driving cycles that mimic onroad conditions under various levels of congestion are being used to develop cycle-based speed correction factors. New certification testing cycles also promise to reduce the frequency of onroad enrichment. The “USO6” cycle represents emissions in aggressive driving, and the “SCO3” cycle reflects the effects of accessory loads like AC use, power steering etc.

A variety of modal modeling approaches are currently under development. A modal emissions model being developed by UC Riverside (An, et al., 1997), is based upon 300 vehicles tested under a variety of laboratory driving cycles. Two different modal approaches developed at the Georgia Institute of Technology are included within the GIS-based modal emission model: an aggregate modal model based upon statistical analysis of historic laboratory data (Guensler et al., 1998) and a load-based prediction module based upon analysis of instrumented vehicle data (Rodgers et al., 1999).

For the past five years, the Georgia Tech Research Partnership has been developing a research-grade motor vehicle emissions model within a geographic information system (GIS) framework. The aggregate modal model within MEASURE model predicts emissions from light-duty vehicles as a function of vehicle operating “mode”, representing a spectrum of vehicle operating conditions including cruise, acceleration, deceleration, idle, and the power demand conditions that lead to enrichment, or high fuel:air ratios. The aggregate modal model is not a look-up table for speed and acceleration. In contrast, the model specifically accounts for interactions between vehicle fleet technology characteristics and vehicle operating modes. For each technology group within the light-duty motor vehicle fleet, the relationships between modal activity and emissions can differ significantly. The aggregate modal model emission rates are based upon theoretical engine-emissions relationships that have been modeling using various statistical techniques (Fomunung et al., 1999). The emission rate models have been estimated through a process that utilizes the best aspects of hierarchical tree-based regression (HTBR) (Breiman, et al., 1995), and ordinary least squares (OLS) regression. The relationships are dependent on both modal and vehicle technology variables. The relationships are “aggregate”, because they rely on ‘bag’ data to derive their modal activities (Washington, 1994), and thus are suitable with existing aggregate approaches contained in the TDM framework.

Emission rates are developed for vehicle technology groups that represent different technology characteristics that have been shown to respond differently to vehicle load and driver input through statistical analysis. Vehicles within a technology group perform similarly with respect to both mean emissions under the FTP and mean emissions response with respect to changes in vehicle operating modes. Emission rate algorithms within the MEASURE predict emission rates for technology groups as a function of the conditions under which the vehicles are operating, specifically various aggregate measures of theirs peed and acceleration profiles. Under a given set of operating conditions, some vehicle technology groups respond with significant emissions increases, while other vehicle technology groups do not. For example, vehicles with relatively small engines and low available power tend to go into fuel enrichment under low to moderate vehicle loads relative to vehicles with larger engines. Vehicle activity variables modeled in MEASURE include average speeds, acceleration rates, deceleration rates, idle time, and surrogates for power demand imposed on the engine. Vehicle and emission control technology variables include fuel metering system, catalytic converter type, supplemental air injection, and transmission speed.

The MEASURE model also tracks the activity of discrete vehicle technology groups in the onroad fleet so that the activity can be appropriately linked with modeled emission rates. The MEASURE framework allows for facility-level aggregations of microscopic traffic simulation, or disaggregation of traditional macroscopic four-step travel demand forecasting models to develop emission-specific vehicle activity data. Once validation and peer review efforts are complete, the model will be submitted to EPA for approval for use in conformity analyses and air quality planning.

Although previous MEASURE-related publications have detailed model estimation methods, to date the authors have not reported the results of model validation. Previous publications have also never offered a comparison of the MEASURE model emission rate algorithms to those employed in the USEPA's MOBILE5a emission rate model. Model validation, using a sample of ‘external’ data to assess model predictive abilities, is perhaps the single most important measure of a model’s ability to capture relationships across space and across time. In addition, it is the only way in which to fairly compare two competing models.

This paper provides an external validation effort for the aggregate modal model option currently included in the MEASURE model. In 1998 and 1999, the USEPA collected emissions test data from 50 technologically diverse vehicles, tested across as many as 16 different dynamometer cycles representing a range of onroad vehicle operations. The USEPA collected the independent data set of 798 dynamometer test results for use in developing the MOBILE6 model. Because neither MOBILE5a nor MEASURE employed these test results during model estimation, and because these vehicles were tested on different driving cycles than were used to estimate the MEASURE and MOBILE5a models, the tests serve as an ideal external model validation sample for assessing and comparing the performance of the hot-stabilized exhaust algorithms in the two models. The performance of the MEASURE and MOBILE are compared using mean absolute prediction errors, linear correlation coefficients between observed and predicted emissions, and mean absolute bias. Results are provided for each driving cycle and for vehicle technology classes. The statistical comparisons clearly indicate that the MEASURE model provides significant improvements in both average emissions estimates and explanatory power over MOBILE5a for all three pollutants across almost every operating cycle tested. The most significant improvements arise in the CO and HC estimates, but even NOx emissions, which are a function of average speed, improve under the MEASURE modeling regime.

MEASURE AGGREGATE MODAL MODEL

The model estimation data consisted of more than 13,000 laboratory tests conducted by the EPA and CARB using standardized test cycle conditions as well as alternative driving cycles (Fomunung, et al, 1999). The aggregate modal model algorithms presented below were estimated using a response variable as the logarithm of the emission rate ratio for each pollutant (Fomunung, et al., 1999). The ratio is the vehicle emission rate (in grams/second) driven on a given cycle (or equivalently across a specified speed/acceleration matrix), divided by that vehicle's emission rate (grams/second) obtained from the FTP Bag 2 testing cycle. The MEASURE Aggregate Modal Model predicts the ratio of g/second emission rates for several vehicle technology groups. The following sequence of equations shows the method for calculating the predicted emission rates for each pollutant in units of either g/second or g/mile:

$$\Psi_i \text{ (g/sec)} = \Psi_i \text{ (g/mile)} * \text{DIST (miles)} / \text{DUR (sec)} \tag{1}$$

$$\Psi_{iBag2} \text{ (g/sec)} = \Psi_{iBag2} \text{ (g/mile)} * 3.91/866 \tag{2}$$

$$R_i \text{ (rate ratio)} = P_i \text{ (g/sec)} / \Psi_{iBag2} \text{ (g/sec)} \tag{3}$$

In these equations, Ψ_i is the observed or measured pollutant (*i* is the index for CO, HC or NOx), P_i is the predicted value of pollutant *i*, Ψ_{iBag2} is the observed FTP Bag2 rate for pollutant *i* for a given vehicle, DIST is the driving cycle distance in miles, DUR is the cycle duration in seconds, 3.91 is the hot stabilized FTP Bag 2 sub-cycle distance in miles, and 866 is the FTP Bag2 sub-cycle duration in seconds.

The MEASURE emissions models are presented in two formats: first they are presented in the form in which they were estimated (suitable for making statistical inferences), and second, they are presented in original variable units, which is often more intuitive for use in emission rate prediction and for interpretation of results. The statistical details of the models are provided in Tables 1, 2, and 3 for CO, HC, and NOx, respectively.

MEASURE Aggregate Modal Model Functional Form (Estimation)

Equation (4) shows the estimation form for CO, Equation (5) shows the estimation form for HC, and Equation (6) shows the estimation form of NOx;.

For CO,

$$\begin{aligned} \text{Log}R_{CO} &= 0.0809 + 0.002*\text{AVGSPD} + 0.0461*\text{ACC.3} + 0.0165*\text{IPS.60} - \\ &0.0283*\text{ips45sar2} + 0.3778*\text{ips90tran1} - 0.0055*\text{tran3idle} + 0.1345*\text{tran5km1} + \\ &0.3966*\text{finj3sar3} - 0.0887*\text{cat3tran1} - 0.2636*\text{sar3tran4} - 0.481*\text{flagco} \end{aligned} \tag{4}$$

Where,

AVGSPD	is the average speed of the driving cycle in mph,
ACC.3	is the proportion of the driving cycle on acceleration greater than 3mph/sec
IPS.X	is the proportion of the driving cycle on inertial power surrogate (IPS) (speed x acceleration) greater than X mph^2/sec (Washington, 1994).Thus IPS.60 implies IPS greater than 60 mph^2/sec.
ips45sar2	is an interaction between IPS.45 and a vehicle with no air injection
ips90tran1	is an interaction variable for a vehicle with automatic transmission on IPS.90
cat3idle	is an interaction variable for a 3-speed manual transmission at idle

tran5km1	is an interaction variable for a 5-speed manual transmission vehicle with mileage <= 25k miles
finj3sar3	is an interaction variable for a vehicle that has throttle body fuel injection and pump air injection,
cat3tran1	is an interaction variable for a vehicle with automatic transmission and TWC,
sar3tran4	is an interaction variable for a vehicle with 4-speed manual transmission and pump air injection
flagco	is a flag used to tag a high emitting vehicle under CO emissions (Wolf et al, 1998).

For HC,

$$\begin{aligned}\mathrm{LogR}_{HC} = {} & 0.0451 - 0.6707*my79 - 0.1356*my82 + 0.0019*AVGSPD + \\ & 0.2021*finj2tran4 + 0.1795*cat2sar1 + 0.1651*cat3sar1 + 0.0318*cat3sar2 - \\ & 0.1189*sar3tran1 + 0.5646*sar1tran5 + 0.0004*cid - 0.2581*sar3km1 - \\ & 0.0169*finj2km3 - 0.5144*flaghc - 0.0129*acc1finj2 - 0.1626*acc3cat2 - \\ & 0.3891*ips90sar3 + 0.0307*dps8finj2 \end{aligned} \tag{5}$$

Where,

my79	is model year < 79
my82	is 79 < model year < 83
finj2tran4	is an interaction variable for a 4-speed manual transmission vehicle with a carburetor,
cat2sar1	is a variable for a pre 1981 model year vehicle with ‘oxidation only’ catalyst and of unspecified air injection type,
cat3sar1	is a variable for a pre 1981 model year vehicle with a TWC and of unspecified air injection type,
cat3sar2	is a variable for a vehicle with TWC and no air injection,
sar3tran1	is an automatic transmission vehicle with pump air injection,
sar1tran5	is a pre 1981 model year, 5-speed manual transmission vehicle of unspecified air injection type,
cid	cubic inch displacement,
sar3km1	is a vehicle with pump air injection and mileage <=25k miles,
finj2km3	is a vehicle with pump air injection and 50k < mileage <= 100k miles,
flaghc	is a high emitting vehicle flag under HC emissions,
acc1finj2	is a carburetor- equipped vehicle operating with acceleration greater that 1mph/sec,
acc3cat2	is an ‘oxidation only’ catalyst vehicle on ACC.3,
ips90sar3	is a vehicle with air pump on IPS.90,
dps8finj2	is the proportion of drag power surrogate (DPS) (speed x speed x acceleration) greater than 8 mph^3/sec.

For NOx,

$$LogR_{NOx} = -0.5864 + 0.0225*AVGSPD + 0.3424*IPS.120 + 0.6329*ACC.6 + 0.0247*DEC.2 + 0.0083*finj2km1 + 0.0028*finj2km2 - 0.0021*cat2km3 + 0.0026*cat3km2 + 0.0003*cat3km3 - 0.0085*finj1km3flagnox - 0.0068*finj3km3flagnox \quad (6)$$

Where,

IPS.120 is IPS > 120 mph^2/sec,
ACC.6 is proportion of acceleration greater than 6 mph/sec,
DEC.2 is proportion of deceleration greater than 2 mph/sec,
finj2km1 is a carburetor equipped vehicle with mileage < 25k miles,
finj2km2 is a carburetor equipped vehicle with 25k < mileage <= 50k miles,
cat2km3 is an 'oxidation only' catalyst vehicle with 50k < mileage <= 100k miles,
cat3km2 is a TWC vehicle with 25k < mileage <= 50k miles,
cat3km3 is a TWC vehicle with 50k < mileage <= 100k miles,
finj1km3flagnox is a second order interaction variable for a high emitting vehicle with port fuel injection and 50k < mileage <= 100k miles,
finj3km3flagnox is a second order interaction variable for a high emitting vehicle with throttle body fuel injection and 50k < mileage <= 100k miles.

On a vehicle by vehicle basis, this implies that after calculating R_i from the response variable, the predicted rate in g/second is:

$$P_i\ (g/sec) = R_i * \Psi_{iBag2} \quad (7)$$

Note, Equation 7 is similar in form to the embedded algorithm in MOBILE which gives emission rates as Correction Factors * BER. Where BER stands for base emission rate, akin to Ψ_{iBag2}; R_i represents all the variables which feature in the models for each pollutant and can be thought of as speed, load, and technology correction factors. The conversion to g/mile is straight forward,

$$P_i\ (g/mile) = R_i * \Psi_{iBag2} * 1/AVGSPD \quad (8)$$

MEASURE Emission Model Functional Form (Prediction)

The prediction forms for CO, HC, and NOx, are shown in Equations (9), (10), and (11), respectively, and the variables are as previously described.

For CO,

$$P_{CO}\ (g/sec) = 1.205 * FTP\ Bag2 * antilog\{0.0809 + 0.002*AVGSPD + 0.0461*ACC.3 + 0.0165*IPS.60 - 0.0283*ips45sar2 + 0.3778*ips90tr1 - 0.0055*tran3idle + 0.1345*tran51 + 0.3966*finj3sar3 - 0.0887*cat3tran1 - 0.2636*sar3tran4 - 0.481*flagco\} \quad (9)$$

For HC,

$$P_{HC}\text{ (g/sec)} = 1.11 * \text{FTP Bag2} * \text{antilog}\{-0.6707*\text{myhc28} - 0.1356*\text{myhc81} + 0.0019*\text{AVGSPD} + 0.2021*\text{finj2tran4} + 0.1795*\text{cat2sar1} + 0.1651*\text{cat3sar1} + 0.0318*\text{cat3sar2} - 0.1189*\text{sar3tran1} + 0.5646*\text{sar1tran5} + 0.0004*\text{cid} - 0.2581*\text{sar31} - 0.0169*\text{finj23} - 0.5144*\text{flaghc} - 0.0129*\text{acc1finj2} - 0.1626*\text{acc3cat2} - 0.3891*\text{ips90sar3} + 0.0307*\text{dps8finj2}\} \quad (10)$$

For NOx,

$$P_{NOx}\text{ (g/sec)} = 0.259 * \text{FTP Bag2} * \text{antilog}\{0.0225(\text{AVGSPD}) + 0.3424(\text{IPS .120}) + 0.6329(\text{ACC.6}) + 0.0247(\text{DEC.2}) + 0.0083(\text{finj2km1}) + 0.0028(\text{finj2km2}) - 0.0021(\text{cat2km3}) + 0.0026(\text{cat3km2}) + 0.0003(\text{cat3km3}) - 0.0085(\text{finj1km3flagnox}) - 0.0068(\text{finj3km3flagnox})\} \quad (11)$$

VALIDATION DATA SET DESCRIPTION

Two types of modal validation efforts are typically performed: internal and external. Internal validation consists of model checking for plausibility of signs and magnitudes of estimated coefficients, agreement with past models and theory, and model diagnostic checks such as distribution of error terms, normality of error terms, etc. Internal validation was performed as part of the model estimation procedure, and is documented in Fomunung, et al. (1999). External validation is the process whereby a model is compared to data collected ‘outside’ the modeling framework, i.e., data from another location or time. External validation is the only way to check if a model has been ‘overfit’ to data, thus capturing spurious rather than real relationships or underlying structure in the data. It is also the only way to determine whether the relationships captured in the estimated model reflect the same relationships elsewhere, or over time. Finally, it is the only objective way to compare and contrast two models estimated on different data. The desire to assess MEASURE transferability and to compare MEASURE performance to that of the MOBILE5a emission prediction model have motivated these external validation efforts.

The data used for MEASURE and MOBILE5a validation consists of 50 vehicles tested across 16 different hot-stabilized driving cycles. Neither the MEASURE nor MOBILE5a models were estimated using these data. Table 4 lists these different cycles, and shows the average speed, maximum speed, and acceleration characteristics for each cycle. The EPA tested each vehicle on every cycle (three cycles only included 49 of the 50 vehicles). The balanced sampling design results in vehicle to vehicle, within vehicle, and cycle to cycle variability. A similar table for the MEASURE model development data set is shown in Table 5. There are a number of differences between the two data sets: only two driving cycles, NYCC, and the CARB “unified” cycle were used in both data sets; average speeds range from 2.45 mph to 59.9 mph in the model data, and from 7.1 to 63.2 mph in the validation data; maximum speeds range from 10 to 71.3 mph in the model development data, and from 27.7 to 74.7 mph in the validation data; maximum acceleration range from 1.5 to 6.9 mph/sec in the model development data, and from 2.3 to 6.9 mph/sec in the validation data. These differences notwithstanding, the independence of the validation data set lends itself well for purposes of model evaluation.

The in-use data obtained from past testing efforts by the both the USEPA and the CARB and used to develop the MEASURE aggregate modal emission models are actually inferior to the validation data in several respects. First, the data used to develop MEASURE did not include all vehicles tested on all cycles. Second, the vehicles recruited, in aggregate, are not representative of the national on-road vehicle fleet. Finally, there is little to no replicate testing, so within-driver variability is unknown. However, the shear size of the aggregate database available for estimating the MEASURE model emissions algorithms reconciles to some extent some of these potential shortcomings.

PREPARING THE MEASURE AND MOBILE MODELS FOR VALIDATION AND COMPARISON

Prior to assessing the predictive abilities of both MEASURE and MOBILE5a, it was necessary to set some ground rules for model validation and comparison. First, it was necessary to determine which 'classes' of vehicles would be compared. In other words, it seemed that for at least some comparisons it would be useful to see how the two models predict emission rates for classes of vehicles that are fundamentally different, since emissions are characteristically different for different classes of vehicles. Then, emission rates had to be converted to comparable and meaningful units (gram/second emission rates). Finally, model goodness of fit criteria needed to be established.

Technology Classes

Four different emission control technologies were investigated during model development: fuel injection type, catalytic converter type, transmission speed type, and the supplemental air injection type. Each discrete variable has several levels represented by the abbreviated codes indicated below:

- Fuel Injection (FINJ)
  1. Port fuel injection (PFI), coded as finj1
  2. Carburetor and all pre-1980 domestic cars, coded as finj2
  3. Throttle body fuel injection(TBI), coded as finj3
  4. Unspecified type pre-1980 import and both 1980 domestic and import, coded as finj4
- Catalytic Converter (CAT)
  1. None, coded as cat1
  2. Oxidation only, coded as cat2
  3. 3-way catalyst, coded as cat3
  4. Oxidation and 3-way catalyst, coded as cat4
- Supplemental Air Recirculation (SAR)
  1. Pre-1980 of unspecified type, coded as sar1
  2. None, coded as sar2
  3. Air pump, coded as sar3
  4. Pulse, coded as sar4
- Transmission Speed (TRAN)
  1. Automatic, coded as tran1
  2. Semi automatic, coded as tran2

3. 3-speed manual, coded as tran3
4. 4-speed manual, coded as tran4
5. 5-speed manual, coded as tran5

To capture the effects of deterioration, accrued test vehicle mileage was used as a surrogate for deterioration. Fomunung, et al. (1999) previously determined that deterioration could be best modeled as a step function rather than a linear deterioration rate over time. Thus, four deterioration mileage groups (or bins) were employed in the models. These groups are “<=25,000 miles”, “25,000 <mileage <= 50,000 miles”, “50,000 < mileage <=100,000 miles”, and “mileage >=100,000 miles”, and are labeled km1, km2, km3, and km4, respectively.

It was a fairly complex task to implement the regression equations inside the MEASURE model. First, it was necessary to define mutually exclusive technology groups that would interact uniquely with vehicle operating modes. In essence, it was necessary to employ classification rules that resulted in mutually exclusive and collectively exhaustive technology groups. To define preliminary classification rules, a matrix of all possible combinations of the four technology variables plus the mileage bins and high-emitter status that appear in the regression model (a total of 4 x 4 x 4 x 5 x 4 x 2 or 2560 technology rules) was created for each pollutant. Then using Equation 3 and each of Equations 9, 10, and 11 (which include technology and modal interactions) for CO, HC, and NOx, respectively, the predicted emission rate ratio for each pollutant was computed for each of the 2560 initial classification rules using the modal variables from the highway fuel economy test. Classification rules that yielded the same predicted emission rate ratios for any given cycle were then clustered together (i.e. collapsed into the mutually exclusive technology groups that are represented in the regression equation). A cross-check with modal variables from other driving cycles (LA4, Low Speed 1 and High Speed Cycle 1) produced the same technology groups. Each technology group cluster was then assigned an aggregate definition to represent a “technology group”, as distinct from the former “classification rule”. Consequently, 44 mutually exclusive technology groups were defined for CO, 120 for HC, and 13 for NOx. The vehicle activity of each of these technology groups is then tracked separately within the MEASURE model because the technology and modal activity interaction variables that appear in the regression equations are different for each group.

Emission Rates

The next step was to predict emissions for each pollutant for any given cycle and technology group. To predict emissions for each technology class, one at a time, Equation 7 is modified to the form,

$$P_i \text{ (g/sec)} = R_{ij} * \Psi_{ijBag2} \qquad (12)$$

Where Ψ_{ijBag2} is now defined as the average of the base emission rate (FTP Bag 2) in g/second, of pollutant *i* for technology class *j*. Note again, that in this form the term R_{ij} in Equation 12 represents a cycle-specific correction factor for each technology class. R_{ij} is the predicted rate ratio of pollutant *i*, for technology class *j*. The values for Ψ_{ijBag2} are obtained from the FTP Bag2 measurements in the original date set, while values for R_{ij} depend on the modal variables inputted in the model.

Establishing Model Goodness-of-fit Criteria for Model Validation and Comparison

There are a number of model goodness of fit criteria that can be used to assess the difference between predicted (by MEASURE and MOBILE5a) and observed emissions (the validation data). It should be noted, however, that the focus here is on point estimates of emissions and not precision of estimates. In other words, models are not compared based on the confidence in mean emission predictions, despite the fact that such comparisons would be extremely useful. We omit these comparisons for two important reasons. First, the data set used to estimate the emissions models within MEASURE is much larger than that used for estimating MOBILE5a, and thus statistical estimates are likely to be more precise for MEASURE. Second, the regulatory arena in which models are employed has yet to embrace the use of confidence intervals on model outputs, thus, comparisons of model efficiency would not likely lead to a strong argument for one model over another, since precision is not applied in practice. It is not without objection that we have left these comparisons out, we strongly believe that these types of comparisons are valid criteria for mounting evidence in favor of one model over another.

Instead, the focus in this validation exercise is on point estimates of emissions. Because models are used to provide point estimates of emission rates (by technology class, etc.), then the comparison between predicted and actual emissions is by far the most relevant model validation exercise to undertake. In addition, overall model bias(the mean difference between predicted and observed emissions for a sufficiently large validation sample) reflects perhaps the most important criterion for comparing whether a model is working well in practice.

The relative performance of the two models, MEASURE and MOBILE5a, is based on three statistical measures of effectiveness: the linear correlation coefficient (Ott, 1993), the root mean squared prediction error (RMSPE) (Neter et al, 1996), and the mean prediction bias. The linear correlation coefficient reflects the degree of probability that a linear relation exists between observed and predicted emissions. A high linear correlation coefficient would imply a good degree of linearity between paired data (predicted and observed emissions for vehicle *i*), whereas, a low coefficient would imply the reverse. The RMSPE is a measure of the error component of prediction. When comparing two models, the model with a smaller value of RMSPE is a better predictor of the observed phenomenon. In and of itself, a high linear correlation coefficient may not be a sufficient determinant of how well a model predicts, because in practice, it is possible to have high biases (or errors) and still have a high correlation coefficient. However, when considered together, low values of RMSPE in concert with a high linear correlation coefficient is a good indicator that a model predicts well. The third measure of comparison is mean prediction bias. Prediction bias measures the systematic under- or over-prediction of a model over a set of predictions. When considering forecasting with a model, prediction bias is very undesirable because it leads to a systematic under- or over-prediction of the quantity being predicted. Ideally, mean prediction bias is arbitrarily close to zero, so that prediction of observations, on average, is correct.

The MOBILE5a hot stabilized emission rates for each vehicle in the data set were predicted from the FTP Bag 2 hot stabilized emission rate for each vehicle. The researchers modified the MOBILE5a input file provided by the USEPA Region 4 office for Atlanta to reflect 100% hot stabilized operations (by setting the fractions of cold and hot start VMT to zero). The model was then set in a model year mode to predict emission rates for each model year. The model was then run in 5-mph average speed increments to develop an emission rate matrix by model year and average speed for calendar year 1997. A matrix of emission rate ratios was developed from the emission rate matrix, with the 20-mph emission rate serving as the baseline emission rate (to conform with MOBILE5a internal assumptions related to the 19.6-mph average speed of the composite FTP). The emission rate ratio is equivalent to the speed correction factor implemented by MOBILE5a for each model year. The emission rate ratio for the average speed of the test cycle (found in the matrix using cubic spline interpolation) was then multiplied by the hot stabilized FTP Bag 2 emission rate for that vehicle to estimate emissions on the alternative test cycle.

RESULTS OF MODEL VALIDATION EXERCISE

This section describes the results of the validation of the MEASURE and MOBILE5a emission factor modules by comparing their prediction abilities across the set of validation data. Using validation vehicle characteristics and emissions results for each of the three pollutants CO, HC, and NOx, the MEASURE and MOBILE5a emissions algorithms (shown in Equations 9, 10, and 11 respectively for the MEASURE model algorithms) were used to predict the observed data. The performance of MEASURE and MOBILE5a and the three validation criterion are provided in the following tables.

Because a vehicle fleet is usually tracked (in practice) by characterizing the number of vehicles in each technology class and by model year, model validation results were computed both for aggregate data (all vehicles) by driving cycle and by technology class. The results provided on a driving cycle basis provides information on how well the models explain variability in emissions due to differences in modal activity or driving profiles, while technology-class based results provide information on how well the models explain emissions differences caused by differences in vehicle technologies.

The results of the performance evaluation are shown in Tables 6 through 11. The linear correlation results on a cycle basis are shown in Table 6, while Tables 7 a, b, and c list the results on a technology class basis for CO, HC, and NOx respectively. The number of vehicles tested on each cycle was shown in Table 4, whereas Tables 7a, b, and c, shows that the total 797 vehicle-tests in the validation data set were distributed into 16 CO technology classes, 13 HC technology classes, and 5 NOx technology classes. In addition, the sample size of each technology class is shown. The results in Table 6 show that for CO and HC, the MEASURE model out performs MOBILE across all test cycles (the highest linear correlation coefficient in each comparison is bolded) , while for NOx both models perform equally across almost all cycles with MEASURE doing better in the rest of the cycles. For the CO and HC results in Table 7, no general trend is discernible, but it can be noted that for a majority of the results MEASURE performs equally well or better than MOBILE; for NOx however, MOBILE performs slightly better than MEASURE in 4 technology classes, and significantly better in technology class 7.

In Tables 8 and 9 are shown the results from the root mean square prediction error analysis (the smallest RMSPE is bolded in each comparison). Table 8 shows the results on cycle basis, and Table 9 on technology class basis. As with the case of linear correlation coefficient, the results on a cycle basis indicate that for CO and HC, MEASURE performs better than MOBILE, but for NOx MEASURE performs equally well or slightly better than MOBILE. On a technology class basis, MEASURE is only marginally better than MOBILE for CO and HC, and results are mixed for NOx.

Table 10 shows the result of the mean prediction bias on a cycle basis, and Table 11 shows the results on technology group basis (smallest absolute mean prediction bias is bolded in each comparison). Both tables also provide overall weighted average mean prediction bias per pollutant. When comparing mean prediction bias, it can be seen that MEASURE consistently under-predicts , while MOBILE consistently over-predicts, both on a cycle and technology class bases. However, the same results indicate that across all cycles and technology classes the degree of under-prediction (as measured by the magnitude of the biases) by MEASURE is lower than that of over-prediction by MOBILE, demonstrating once again by this measure of assessment that MEASURE performs better than MOBILE.

CONCLUSIONS

The MEASURE model showed consistent improvements in correlation between observed and predicted emissions for the validation data set The consistent improvement in correlation coefficient suggests that the additional explanatory variables and their interactions employed in the MEASURE model provide additional explanatory power. The relatively smaller improvement in NOx predictions stems from the fact that the average speed approach to modeling NOx emissions is not significantly inferior to using improved vehicle activity information…average speed works quite well for NOx.

Some of the driving cycles used in the validation study were designed by EPA contractors to represent onroad driving conditions under varying levels of congestion. Many of these cycles are significantly different from those that were used to develop the MOBILE5a and MEASURE models. The strong performance of the MEASURE model on these new cycles reveals the strength of applying the model to cycles outside those used to develop the model. These findings provide empirical support for the underlying principle that although the models are cycle-based and aggregate, the discrete contributions of various modal contributions have been modeled well in the MEASURE modeling algorithms, and can be used to model the emissions resulting from a variety of “off-cycle” vehicle activities.

In general the results provided here are encouraging for MEASURE. The general superiority of MEASURE on mean prediction bias suggests if MEASURE and MOBILE5a were applied in practice for forecasting, MEASURE predictions would be more accurate, on average, by a factor of 1.8. On the basis of each pollutant, MEASURE would be more accurate by a factor of 1.6 for CO, 1.7 for HC, and 1.4 for NOx. This is an extremely compelling reason to favor MEASURE over MOBILE5a, since systematic errors in emission rates are multiplied by the number of vehicles in an urban area, and then again by the amount of mileage driven on a ‘typical day.’ MEASURE does slightly under-predicts emissions for these vehicles, but this is not a huge concern. MEASURE would also slightly-under predict emissions reductions likely to be garnered from proposed control strategies, so there is no major impact associated with using the model for control strategy modeling (i.e. as a comparative tool across control strategies and time). Furthermore, the data used to develop MEASURE contained very few test results from 1994 and later model year vehicles. When new data from laboratory studies (such as the UC Riverside study) are included in the data set and the MEASURE algorithms are re-derived, the authors expect further improved performance in applications to the modern vehicle fleet.

Additional validation exercises of MEASURE are currently underway in Atlanta, Georgia. A vertical pollutant flux study was coordinated with detailed monitoring of a metered ramp system on the I-75, north of downtown Atlanta. The research team collected traffic flow data, vehicle classification, fleet technology characteristics (by monitored license plate data and later decoding the registration VINs), and speed acceleration profiles with laser guns on the 4 ramps and the proximal freeway links. The research team is currently analyzing the 18 days of vehicle activity data and will use the vertical pollutant flux results to compare predicted and measured emission rates as a means for comparing real world to modeled emissions

REFERENCES

An, F., M. Barth, J. Norbeck, and M. Ross. “The Development of a Comprehensive Modal Emissions Model: Operating Under Hot-Stabilized Conditions”. In Transportation Research Record 1587, TRB, National Research Council, Washington, D.C., 1997, pp.52-62.

Bishop, G.A.; Stedman, D. H.; and Ashbaugh, L. “Motor Vehicle Emissions Variability”, J. Air & Waste Manage. Assoc., 1996, 46(7):667-675.

Brieman, Leo, Friedman, J.H., Olshen, R.A., Stone C.J., Classification and Regression Trees. Wardsworth International Group, Belmont, CA., 1984

Cadle, S. H., R. A. Gorse, and T. C. Belian. “Real-world vehicle Emissions: A Summary of the Seventh CRC On-road Vehicle Emissions Workshop”, San Diego, CA April 9-11, 1997

Cicero-Fernandez, Pablo, Jeffrey Long, Arthur Winer, "Effects of Grades and other Loads on On-Road Emissions of Hydrocarbons and Carbon Monoxide". J. Air & Waste Manage. Assoc., Pittsburgh, PA., Vol. 47, 1997.

Cicero-Fernandez, Pablo, Walter Wong, Jeffrey Long, "Fixed Point Mobile Source Emissions Due to Terrain Related Effects: A Preliminary Assessment". Proceedings of the 90th Annual Meeting, AWMA, Pittsburgh, PA., June, 1997.

Fomunung, I , S. Washington, R. Guensler. "A Statistical Model for Estimating Oxides of Nitrogen Emissions from Light Duty Motor Vehicles". Transportation Research D: Transport and Environment 4/5, 1999, pp. 333-352

Gertler, A.W., J. C. Sagiebiel, D.N. Wittorf, W. R. Pierson, W. A., Dippel, D Freeman, L. Sheetz, "Vehicle Emissions in Five Urban Tunnels" CRC Project No. E-5 Final Report. March 1997.

Guensler, Randall, "Data Needs for Evolving Motor Vehicle Emission Modeling Approaches". In Transportation Planning and Air Quality II; Paul Benson, Ed.; American Society of Civil Engineers, New York, NY., 1993

Guensler, Randall, Simon Washington, and William Bachman, "Overview of the MEASURE Modeling Framework". In Transportation Planning and Air Quality III, Emerging Strategies and Working Solutions, Simon Washington ED., American Society of Civil Engineers, 1998, pp. 51-70

Jimenez-Palacio J.L. (1999), "Understanding and Quantifying Motor Vehicle Emissions with Vehicle Specific Power and TILDAS Remote Sensing", Dissertation Report, Massachusetts Institute of Technology

Neter, J., Kutner, M. H., Nachtsheim, C. J., Wasserman, W., Applied Linear Statistical Models, 4th Edition, IRWIN, Boston, MA., 1996

Ott, R. Lyman, An Introduction to Statistical Methods and Data Analysis 4th Edition, Duxbury Press, Wardsworth Publishing Company, Belmont CA., 1993

Pierson, W.R., A. W. Gertler and R. L. Bradow. "Comparison of the SCAQS Tunnel Study with other On-Road Vehicle Emission Data". Journal of the Air and Waste Management Association 40: 1495-1504, 1990.

Rodgers, M.O., 1995. "An Improved Power Demand Model for Evaluation of On-Road Vehicle Emissions"; The Emission Inventory: Programs and Progress (EPA/AWMA); October 95.

Shih, R.; Fable, S.; Sawyer, R.F. "Effects of Driving Behavior on Automotive Emissions"; Proceedings of the Seventh CRC On-Road Vehicle Emissions Workshop, Coordinating Research Council, Atlanta, GA, pp. 6-115 to 6-124, 1997.

Washington, S. “Estimation of a Vehicular Carbon Monoxide Modal Emissions Model and Assessment of an Intelligent Transportation Technology”. Dissertation. Institute of Transportation Studies, University of California at Davis, 1994.

Webster, W.J.; Shih, C. “A Statistically-Derived Metric to Monitor Time-Speed Variability in Practical Emissions Testing”; Proceedings of the Sixth CRC On-Road Vehicle Emissions Workshop, held March 18-28, 1996 in San Diego, CA, published by Coordinating Research Council, Atlanta, GA.

Wolf, J, R. Guensler, S. Washington, and W. Bachman, “High-Emitting Vehicle Characteristics Using Regression Tree Analysis”. In Transportation Research Record 1641, TRB, National Research Council, Washington, D.C., 1998, pp. 58-65.

Table 1. CO Model Details

Variable	Estimated Coefficient	Std. Error	t-value	Pr(>|t|)
Intercept	0.0809	0.0154	5.2382	<0.001
AVGSPD	0.0020	0.0004	5.0514	<0.001
ACC.3	0.0461	0.0026	17.8998	<0.001
IPS.60	0.0165	0.0066	2.4909	<0.013
ips45sar2	-0.0283	0.0067	-4.2136	<0.001
ips90tran1	0.3778	0.0265	14.2899	<0.001
cat3idle	-0.0055	0.0004	-13.8299	<0.001
tran5km1	0.1345	0.0134	10.0067	<0.001
finj3sar3	0.3966	0.0314	12.6305	<0.001
cat3tran1	-0.0887	0.0145	-6.1218	<0.001
sar3tran4	-0.2636	0.1177	-2.2401	<0.025
flagco	-0.4810	0.0290	-16.5777	<0.001

Table 2. HC Model Details

Variable	Estimated Coefficient	Std. Error	t-value	Pr(>|t|)
intercept	0.0451	0.0172	2.6194	<0.009
my79	-0.6707	0.0388	-17.3088	< 0.001
my81	-0.1356	0.0242	-5.6003	< 0.001
AVGSPD	0.0019	0.0002	8.5775	< 0.001
finj2tran4	0.2021	0.0559	3.6133	< 0.0003
cat2sar1	0.1795	0.0477	3.7627	< 0.0002
cat3sar1	0.1651	0.0646	2.5571	< 0.011
cat3sar2	0.0318	0.0127	2.5104	< 0.012
sar3tran1	-0.1189	0.0216	-5.5146	< 0.001
sar1tran5	0.5646	0.1067	5.2927	< 0.001
cid	0.0004	0.0001	4.2940	< 0.001
sar3km1	-0.2518	0.0185	-13.5982	< 0.001
finj2km3	-0.0169	0.0034	-4.9983	< 0.001
flaghc	-0.5144	0.0276	-18.6432	< 0.001
acc1finj2	-0.0129	0.0013	-9.7293	< 0.001
acc3cat2	-1626	0.0220	-7.3874	< 0.001
ips90sar3	-0.3891	0.0610	-6.3818	< 0.001
dps8finj2	0.0307	0.0015	20.9745	< 0.001

Table 3. NOx Model Details

Variable	Estimated Coefficient	Std. Error	t-value	Pr(>|t|)
(intercept)	-0.5864	0.0068	-85.9273	< 0.001
AVGSPD	0.0225	0.0002	131.6271	< 0.001
IPS .120	0.3424	0.0452	7.5684	< 0.001
ACC.6	0.6329	0.1683	3.7595	< 0.002
DEC.2	0.0247	0.0007	34.8026	< 0.001
finj1	0.0083	0.0008	10.4205	< 0.001
finj2	0.0028	0.0004	6.8670	< 0.001
cat1	-0.0021	0.0004	-5.9243	< 0.001
cat2	0.0026	0.0002	13.5707	< 0.001
cat3	0.0003	0.0001	2.9355	< 0.001
flag1	-0.0085	0.0015	-5.7854	< 0.001
flag2	-0.0068	0.0009	-7.4491	< 0.001

Table 4: Number of Tests, Average Speeds, Maximum Speeds, and Maximum Instantaneous Acceleration of Each Test Cycle- Validation Data Set

Test Cycle Description	Name	# of Tests	Average Speed (mph)	Max. Speed (mph)	Max. Accel. (mph/sec)
Arterial LOS A-B Cycle	ARTA	50	24.8	58.9	5.0
Arterial LOS C-D Cycle	ARTC	50	19.2	49.5	5.7
Arterial LOS E-F Cycle	ARTE	50	11.6	39.9	5.8
Hot Running 505	F505	50	25.6	56.3	3.4
New York City Cycle	FNYC	50	7.1	27.7	6.0
Freeway LOS A-C Cycle	FWAC	50	59.7	73.1	3.4
High-Speed Freeway Cycle	FWHS	50	63.2	74.7	2.7
Freeway LOS D Cycle	FWYD	50	52.9	70.6	2.3
Freeway LOS E Cycle	FWYE	50	30.5	63.0	5.3
Freeway LOS F Cycle	FWYF	50	18.6	49.9	6.9
Freeway LOS G Cycle	FWYG	50	13.1	35.7	3.8
CARB “Unified” LA92 Cycle	LA92	49	24.7	67.2	6.9
Local Roadways Cycle	LOCL	50	12.8	38.3	3.7
Freeway Ramp Cycle	RAMP	50	34.6	60.2	5.7
Start Cycle	ST01	49	20.1	41.0	5.1
Area wide Non-Freeway Cycle	WIDE	49	19.4	52.3	6.4

Table 5: Number of Tests, Average Speeds, Maximum Speeds, and Maximum Instantaneous Acceleration of Each Test Cycle- Model Development Data Set

Cycle Name*	# of Tests	Avg. Speed (mph)	Max Speed (mph)	Max Accel. (mph/sec)
Arterial 1	23	14.3	44.9	6.9
Arterial 2	21	24.06	46.3	5.8
Arterial 3	23	34.39	54.9	6.9
CCDH (Bag 2)	58	13.4	29.8	3.0
Cycle 1	21	59.9	71.3	1.5
Cycle 2	23	53.31	68.0	2.0
Cycle 3	22	39.28	68.9	4.6
Cycle 4	20	31.54	61.9	3.3
Cycle 5	23	23.6	56.5	3.9
Cycle 6	21	15.94	40.9	5.0
Cycle 7	23	9.17	39.7	3.1
HFET	6586	48.2	59.9	3.2
LSP 1	813	2.45	10	2.4
LSP 2	814	3.62	14	2.5
LSP 3	815	4.04	16	3.4
NYCC	1218	7.07	27.7	6.0
SC12	1199	11.7	29.1	3.3
SC36	1201	36.5	57.0	6
Unified Cycle (Bag 2)	88	27.4	67.2	6.9
FTP (Bag 2)	All	16.2	34.3	3.3

*

Arterial 1,2,3 denote cycles developed in California for roadway specific testing
CCDH denotes a cycle developed for use by the Colorado Department of Health for high altitude testing
Cycles 1,2,3,4,5,6, and 7 represent high-speed cycles developed in California for roadway facility testing
HFET stands for Highway Fuel Economy Test
LSP1, 2,3 refers to EPA's low speed testing cycles
SC12 and SC36 refer to EPA speed correction factor cycles
Unified Cycle (LA92) refers to a new California laboratory testing cycle providing greater engine loads

Table 6 Linear Correlation Coefficient: Observed vs. MEASURE, and Observed vs. MOBILE, By Cycle

Cycle	Observed Vs. Predicted- MEASURE (g/sec)			Observed Vs. Predicted - MOBILE5a (g/sec)		
	CO	**HC**	**NOx**	**CO**	**HC**	**NOx**
ARTA	**0.559**	**0.702**	**0.391**	0.268	0.243	0.339
ARTC	**0.463**	**0.577**	**0.411**	0.368	0.199	0.269
ARTE	**0.432**	**0.606**	**0.398**	0.314	0.252	0.280
F505	**0.602**	**0.688**	**0.372**	0.266	0.302	0.313
FNYC	**0.399**	**0.601**	**0.446**	0.314	0.263	-0.0008
FWAC	**0.642**	**0.647**	**0.496**	0.221	0.255	0.282
FWHS	**0.634**	**0.686**	0.428	0.229	0.232	**0.519**
FWYD	**0.581**	**0.522**	0.40	0.395	0.373	**0.513**
FWYE	**0.545**	**0.672**	0.428	0.370	0.323	**0.451**
FWYF	**0.464**	**0.598**	**0.465**	0.368	0.225	0.299
FWYG	**0.458**	**0.581**	**0.389**	0.336	0.264	0.265
LA92	**0.579**	**0.630**	**0.424**	0.355	0.230	0.321
LOCL	**0.465**	**0.616**	**0.434**	0.314	0.235	0.260
RAMP	**0.610**	**0.630**	0.361	0.306	0.357	**0.379**
ST01	**0.665**	**0.689**	**0.509**	0.192	0.151	0.205
WIDE	**0.561**	**0.682**	**0.424**	0.323	0.269	0.273

Table 7a. Linear Correlation Coefficient: Observed vs. MEASURE, and Observed vs. MOBILE, By Technology Class - CO

Tech Class	**Sample Size**	**Observed Vs. Predicted- MEASURE (g/sec)**	**Observed Vs. Predicted - MOBILE5a (g/sec)**
3	16	0.512	**0.514**
6	16	**0.781**	0.548
11	32	0.164	**0.533**
14	190	**0.293**	0.120
19	16	**0.626**	0.467
20	32	0.599	**0.635**
21	64	0.578	**0.877**
22	112	**0.501**	0.456
23	176	0.433	**0.476**
27	16	**0.849**	0.765
33	15	**0.975**	0.908
36	32	0.500	**0.535**
39	16	**0.880**	0.809
40	16	**0.952**	0.908
41	32	**0.735**	0.534
42	16	**0.624**	0.439

Table 7b. Linear Correlation Coefficient: Observed vs. MEASURE, and Observed vs. MOBILE, By Technology Class - HC

Tech Class	**Sample Size**	**Observed Vs. Predicted- MEASURE (g/sec)**	**Observed Vs. Predicted - MOBILE5a (g/sec)**
32	16	**0.597**	0.555
34	16	**-0.099**	-0.065
38	16	0.095	**0.108**
51	16	**0.126**	0.115
54	16	-0.452	**-0.459**
77	304	**0.370**	0.092
80	191	0.145	**0.213**
84	79	**-0.110**	-0.042
95	64	**0.296**	0.111
96	16	**0.539**	0.460
97	15	**0.946**	0.915
108	16	0.075	**0.126**
112	32	0.085	**0.306**

Table 7c. Linear Correlation Coefficient: Observed vs. MEASURE, and Observed vs. MOBILE, By Technology Class - NOx

Tech Class	Sample Size	Observed Vs. Predicted- MEASURE (g/sec)	Observed Vs. Predicted - MOBILE5a (g/sec)
4	16	-0.288	**-0.296**
5	161	0.368	**0.449**
6	556	0.452	**0.497**
7	48	0.746	**0.939**
8	16	**0.926**	0.952

Table 8 Root Mean Square Prediction Error: Observed vs. MEASURE, and Observed vs. MOBILE, By Cycle

Cycle	**Observed Vs. Predicted- MEASURE (g/sec)**			**Observed Vs. Predicted - MOBILE5a (g/sec)**		
	CO	HC	NOx	CO	HC	NOx
ARTA	**0.1038**	**0.0069**	**0.0071**	0.1362	0.0084	0.0072
ARTC	**0.0900**	**0.0078**	**0.0064**	0.1149	0.0086	0.0065
ARTE	**0.0763**	**0.0063**	**0.0048**	0.0980	0.0071	0.0050
F505	**0.0990**	**0.0056**	**0.0072**	0.1268	0.0068	0.0075
FNYC	**0.0679**	**0.0056**	**0.0024**	0.0830	0.0064	0.0029
FWAC	**0.2763**	**0.0095**	0.0166	0.3232	0.0110	**0.0162**
FWHS	**0.3271**	**0.0107**	0.0194	0.3615	0.0121	**0.0180**
FWYD	**0.1947**	**0.0084**	**0.0129**	0.2429	0.0099	0.0138
FWYE	**0.1189**	**0.0073**	**0.0074**	0.1529	0.0089	0.0080
FWYF	**0.0787**	**0.0069**	0.0059	0.1029	0.0078	0.0059
FWYG	**0.0658**	**0.0057**	0.0029	0.0784	0.0061	0.0029
LA92	**0.1099**	**0.0073**	**0.0080**	0.1505	0.0089	0.0085
LOCL	**0.0684**	**0.0058**	0.0042	0.0832	0.0064	**0.0041**
RAMP	**0.2153**	**0.0107**	**0.0121**	0.2666	0.0130	0.0128
ST01	**0.2031**	**0.0202**	**0.0092**	0.2564	0.0238	0.0093
WIDE	**0.0923**	**0.0064**	**0.0057**	0.1266	0.0080	0.0058

Table 9 Root Mean Square Prediction Error (RMSPE):Observed vs. MEASURE, and Observed vs. MOBILE, By Technology Class

MEASURE									**MOBILE**								
CO			HC			NOx			CO			HC			NOx		
Tech Class	n*	rmsper	Tech Class	n*	msper	Tech Class	n*	msper	Tech Class	n*	rmsper	Tech Class	n*	rmsper	Tech Class	n*	rmsper
3	16	0.1088	32	16	**1.09E-02**	4	16	**2.5E-03**	3	16	**0.0199**	32	16	1.2E-02	4	16	3.4E-03
6	16	**0.0705**	34	16	**7.41E-03**	5	161	8.4E-03	6	16	0.1038	34	16	7.7E-03	5	161	**7.5E-03**
11	32	0.1532	38	16	8.33E-03	6	556	7.9E-03	11	32	**0.1449**	38	16	**1.1E-03**	6	556	**7.6E-03**
14	190	**0.0756**	51	16	**1.10E-02**	7	48	**9.9E-03**	14	190	0.0794	51	16	1.2E-02	7	48	1.3E-02
19	16	**0.1206**	54	16	1.20E-02	8	16	**1.9E-02**	19	16	0.1294	54	16	1.2E-02	8	16	3.5E-02
20	32	**0.2249**	77	304	**6.28E-03**				20	32	0.2301	77	304	6.4E-03			
21	64	0.0833	80	191	3.80E-03				21	64	**0.0723**	80	191	**3.7E-03**			
22	112	**0.0550**	84	79	5.11E-03				22	112	0.0644	84	79	**5.8E-03**			
23	176	**0.0384**	95	64	**4.68E-03**				23	176	0.0395	95	64	4.8E-03			
27	16	**0.1086**	96	16	**2.14E-02**				27	16	0.1228	96	16	3.8E-02			
33	15	**0.5747**	97	15	**3.82E-02**				33	15	0.9304	97	15	4.2E-02			
36	32	**0.4201**	108	16	**5.49E-03**				36	32	0.5565	108	16	2.5E-02			
39	16	**0.3824**	112	32	**4.33E-02**				39	16	0.5878	112	32	5.8E-02			
40	16	**1.2554**							40	16	1.3031						
41	32	**0.2514**							41	32	0.4798						
42	16	**0.0877**							42	16	0.3380						

n* is the number of cases in each technology class of the validation data set

Table 10 Mean Prediction Bias: Observed vs. MEASURE, and Observed vs. MOBILE, By Cycle

Cycle	Observed Vs. Predicted- MEASURE (g/sec)			Observed Vs. Predicted - MOBILE5a (g/sec)		
	CO	HC	NOx	CO	HC	NOx
ARTA	**-0.0737**	**-0.00415**	-0.00623	0.134371	0.007821	0.006314
ARTC	**-0.0539**	**-0.00454**	-0.00593	0.112833	0.008136	**0.005882**
ARTE	**-0.0366**	**-0.00304**	**-0.00442**	0.095563	0.006779	0.004769
F505	**-0.06372**	**-0.00251**	**-0.00608**	0.12464	0.00619	0.00629
FNYC	**-0.0223**	**-0.00206**	**-0.00211**	0.08146	0.00634	0.00273
FWAC	**-0.2646**	**-0.00699**	**-0.00149**	0.323182	0.01042	0.011025
FWHS	**-0.31357**	**-0.00824**	**0.00073**	0.359501	0.011016	0.010536
FWYD	**-0.17421**	**-0.00536**	**-0.00478**	0.242362	0.00928	0.010209
FWYE	**-0.08906**	**-0.00457**	**-0.00611**	0.151185	0.008353	0.00667
FWYF	**-0.04174**	**-0.00362**	-0.00537	0.100134	0.00722	**0.005242**
FWYG	**-0.01704**	**-0.0019**	-0.00219	0.074849	0.005532	**0.002057**
LA92	**-0.08664**	**-0.00478**	**-0.00725**	0.149526	0.008463	0.00794
LOCL	**-0.0209**	**-0.00223**	-0.00366	0.078601	0.005845	**0.003484**
RAMP	**-0.20053**	**-0.00894**	**-0.01108**	0.2659582	0.012719	0.011994
ST01	**-0.203**	**-0.02018**	**-0.00908**	0.256356	0.023783	0.009169
WIDE	**-0.00386**	**-0.00386**	-0.00521	0.124249	0.007489	**0.005177**
	-0.10786	*-0.00542*	*-0.00501*	*0.16714*	*0.00907*	*0.00684*

Table 11 Mean Prediction Bias: Observed vs. MEASURE, and Observed vs. MOBILE, By Technology Class

MEASURE									MOBILE								
CO			HC			NOx			CO			HC			NOx		
Tech Class	n^*	bias	Tech Class	n^*	bias	Tech Class	n^*	bias	Tech Class	n^*	bias	Tech Class	n^*	bias	Tech Class	n^*	bias
3	16	0.108813	32	16	-0.01093	4	16	0.00086	3	16	0.01876	32	16	0.011738	4	16	-0.0025
6	16	-0.04523	34	16	-0.00722	5	161	-0.00713	6	16	0.102013	34	16	0.007613	5	161	0.00644
11	32	-0.12028	38	16	0.00793	6	556	-0.0044	11	32	0.141114	38	16	0.000741	6	556	0.00589
14	190	-0.05139	51	16	-0.01102	7	48	-0.00574	14	190	0.076822	51	16	0.011833	7	48	0.01293
19	16	-0.11836	54	16	-0.01199	8	16	-0.00849	19	16	0.129363	54	16	0.011975	8	16	0.03496
20	32	-0.22057	77	304	-0.00446				20	32	0.229696	77	304	0.005755			
21	64	0.029156	80	191	-0.00082				21	64	0.07207	80	191	0.002972			
22	112	-0.05018	84	79	0.00005				22	112	0.063723	84	79	0.005399			
23	176	-0.0275	95	64	-0.00312				23	176	0.035519	95	64	0.00445			
27	16	-0.10864	96	16	-0.02138				27	16	0.122764	96	16	0.03756			
33	15	-0.57475	97	15	-0.03824				33	15	0.930368	97	15	0.042235			
36	32	-0.31805	108	16	-0.00087				36	32	0.556478	108	16	0.025163			
39	16	-0.38237	112	32	-0.036				39	16	0.587837	112	32	0.058187			
40	16	-1.25539							40	16	1.30311						
41	32	-0.24356							41	32	0.479786						
42	16	-0.08084							42	16	0.337981						
		-0.10411			*-0.00542*			*-0.00501*			*0.16714*			*0.00907*			*0.00684*

n^* is the number of cases in each technology class of the validation data set