Comparison of MEASURE and MOBILE5a Predictions Using Laboratory Measurements of Vehicle Emission Factors.

Ignatius Fomunung[1,2], Simon Washington[1], and Randall Guensler[1]

Abstract

As part of on-going research at the Georgia Institute of Technology, three algorithms have been recently estimated for predicting mobile source emissions of carbon monoxide, hydrocarbons, and oxides of nitrogen from light duty passenger vehicles. The Georgia Tech Mobile Emissions Assessment System for Urban and Regional Evaluation (MEASURE) is one of several new modal emissions models designed to improve predictions of CO, HC, and NOx for the on-road vehicle fleet.

Three statistical criteria are used to assess the relative predictive performance of the MEASURE and Mobile5a models on an external data set of observed emission factors of CO, HC, and NOx. These 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 under which the emission factors were measured. The most significant improvements arise in the CO and HC estimates, but even NOx emissions, which are largely a function of average speed, are improved under the MEASURE modeling regime. The paper compares and contrasts various prediction performance measures for the two models, and illustrates the importance of such prediction biases with a simple hypothetical scenario.

Introduction

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 on-road vehicle operations.

[1] School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA 30332
[2] Center for Theoretical Studies of Physical Systems, Clark Atlanta University, Atlanta GA 30314

The USEPA collected the independent data set of 798 dynamometer test results for use in developing the MOBILE6 model. Table 1 lists the 16 different cycles, and shows their average speeds, maximum speeds, and acceleration characteristics. 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.

Table 1: 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. Acceln. (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

Because neither MOBILE5a nor MEASURE (Guensler et al, 1998) 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 both the MEASURE and Mobile5a models.

Hot stabilized emission rate algorithms within the MEASURE were estimated from a data set of more than 13000 hot stabilized dynamometer tests. The algorithms predict emission rates for motor vehicle technology groups as a function of the conditions under which the vehicles are operating, specifically various aggregate measures of their speed and acceleration profiles. Vehicle activity related 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 performance of the MEASURE and MOBILE are compared using mean absolute prediction errors, linear correlation coefficients between observed and predicted emissions, and mean prediction bias. Results are provided for each driving cycle and for vehicle technology classes.

Structure of the MEASURE Algorithms

The three algorithms are based on regression equations that were estimated using a combination of parametric and non-parametric statistical methods (Washington et al, 1997, Fomunung et al, 1999). The final algorithms can be written in functional formats as follows:

CO Model:

$\log R_{CO}$ = f(AVGSPD, ACC.3, IPS.60, ips45sar2, ips90tran1, cat3idle, tran5km1, finj3sar3, cat3tran1, sar3tran4, flagco, ε)

Where:

R_{CO}	Emission rate ratio, defined as a vehicle's CO emission rate on any cycle divided by its corresponding FTP Bag2 emission rate.
f	denotes a linear function.
ε	unexplained random noise in the observed emissions.
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).

HC Model:

$\log R_{HC}$ = f(my79, my82, AVGSPD, finj2tran4, cat2sar1, cat3sar1, cat3sar2, sar3tran1, sar1tran5, cid, sar3km1, finj2km3, flaghc, acc1finj2, acc3cat2, ips90sar3, dps8finj2, ε)

Where:

R_{HC}	Emission rate ratio, defined as a vehicle's HC emission rate on any cycle divided by its corresponding FTP Bag2 emission rate.
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.

NOx Model:

$\log R_{NOx}$ = f(AVGSPD, IPS .120, ACC.6, DEC.2, finj2km1, finj2km2, cat2km3, cat3km2, cat3km3, finj1km3flagnox, finj3km3flagnox, ε)

Where:

R_{NOx}	Emission rate ratio, defined as a vehicle's NOx emission rate on any cycle divided by its corresponding FTP Bag2 emission rate.
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.

As can be seen, several significant and seemingly complex variables have been retained in each model. By assessing from a practical perspective, further insight into their structure can be obtained.

CO Variables:

The model variables indicate that the significant modal predictor variables for carbon monoxide are average speed (*AVGSPD*) and the percent of vehicle activity where acceleration exceeds 3.0 mph/s (*ACC.3*), which relate to all vehicles in the fleet. The product of instantaneous speed and acceleration termed inertial power surrogate (IPS), where IPS is greater than or equal to 60 mph^2 /s (*IPS.60*) also affects the emission rate of CO for the entire fleet.

Three other variables are related to both the vehicle's modal activity and a specific characteristic of the fleet. The percent of time spent during vehicle idling is significant for 3-way catalyst equipped vehicles (*cat3idle*). IPS is significant for vehicles with automatic transmissions when IPS is greater than or equal to 90 mph^2/s (*ips90tran1*). For vehicles with no excess air injection, IPS >= 45 mph^2 /s (*ips45sar2*) is a relevant variable.

HC Variables:

The HC model is influenced by many technology variables. The sub fleet of pre 1980 model year vehicles behaves (in terms of its emissions characteristics) differently from other sub fleets. Likewise, 1980 to 1982 model year vehicles (*my82*) behave differently from other model year vehicles.

Interactions between several fleet characteristics are important. When the sub fleet consists of carburetor-equipped, 4-speed manual transmission vehicles, then their emission characteristics are different from others. The significance of the other technology variables in the model, *cat2sar1, cat3sar2, sar3tran1, and sar1tran5* can be similarly justified.

Two variables, *sar3km1* and *finj2km3* demonstrate the significance of deterioration on HC emission rates for vehicles equipped with either pump air injection or carburetor. The degree of deterioration indicated by the mileage bin also seems to be significant in specifying different emission characteristics of the fleet.

For the entire fleet, average speed (*AVGSPD*) is a significant modal variable. Four of the variables are related to both the vehicle's modal activity and a specific characteristic of the fleet. For example, the percent of time where the acceleration is greater than 1.0 mph/sec (*ACC.1*) is significant for carburetor equipped vehicles; the percent of time where the acceleration is greater than 3.0 mph/sec is significant for oxidation - only equipped cars; and the percent of time spent at IPS greater than 90mph^2/sec is relevant for vehicles equipped with air pumps for supplemental air injection. For carburetor -equipped vehicles, drag power surrogate (*DPS*) greater than 800 mph^3/sec is significant.

NOx Variables:

Four modal variables, average speed, percent of time spent with IPS $\geq$ 120 mph^2/sec, percent of driving time with acceleration exceeding 6 mph/sec, and the percent of driving time with deceleration exceeding 2 mph/sec, are significant for the entire fleet.

The rest of the variables are first order and second order interactions between vehicle characteristics and deterioration. The second order variables are *finj2km1, finj2km2, cat2km3, cat3km2, cat3km3* and indicate that the sub fleet of vehicles which constitute these combinations of characteristics behave differently than other vehicles with respect to NOx emissions. High emitting vehicles with high mileage and port fuel injection or throttle body injection are significantly different with respect to NOx than otherwise similar vehicles.

Mobile5a Variables

In contrast, the internal algorithm for predicting hot stabilized emissions in the Mobile series of models is simple in form but usually tedious to implement. It can be represented as follows:

Basic emission factor =f(average speed, model year, ε) x correction factors

The Mobile series of models relies on the Federal Test Procedure (FTP) emission data to provide the basic emission rates, deterioration rates, and tampering rates. Conditions that are different from those in the FTP are handled through various correction factors, which include speed, and temperature correction factors. As shown in the MEASURE algorithms, there is no need for these correction factors (with the exception of temperature), since any such differences in speed, deterioration rates, and tampering rates are handled explicitly during the model estimation process.

Emission Rates

The emission rates produced by the MEASURE model were derived through least squares estimated regression equations. As seen from the structure of the embedded algorithms within MEASURE it is clear that various sub fleets would have emission rate corrections. Thus it was necessary from the outset to be able to characterize all possible unique sub fleets or technology groups in the general fleet. 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. The details involved in developing these technology groups have been described elsewhere (Fomunung et al, 2000). In essence, it was necessary to employ classification rules that resulted in mutually exclusive and collectively exhaustive technology groups. 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.

Using validation vehicle characteristics and emissions results for each of the three pollutants CO, HC, and NOx, the MEASURE and MOBILE5a emissions algorithms (shown previously for the MEASURE model algorithms) were used to predict the observed data.

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 and DISCUSSION

Establishing Model Goodness-of-fit Criteria for Model 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 omitted these types of comparisons, 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 comparative 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.

Results

This section describes the results of the assessment of the MEASURE and MOBILE5a emission factor modules by comparing their prediction abilities across the set of independent data. 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 using the linear correlation coefficient are shown in Tables 2 through 5. The linear correlation results on a cycle basis are shown in Table 2, while Tables 3, 4, and 5 list the results on a technology class basis for CO, HC, and NOx respectively.

Table 2 Linear Correlation Coefficient: Observed vs. MEASURE, and Observed vs. MOBILE, By Cycle

Cycle	Observed Vs. Predicted-			Observed Vs. Predicted - MOBILE 5a		
	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

Tables 3, 4, and 5 also show that the independent data set of 798 vehicle tests 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 2 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 Tables 3 and 4, 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.

Table 3. 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 - MOBILE 5a (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 4. 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 - MOBILE 5a (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 5. Linear Correlation Coefficient: Observed vs. MEASURE, and Observed vs. MOBILE5a, By Technology Class - NOx

Tech Class	Sample Size	Observed Vs. Predicted- MEASURE (g/sec)	Observed Vs. Predicted - MOBILE 5a (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

Unlike for linear correlation coefficient, it is more intuitive to present the results from the other two criterion through a series of charts as depicted in Figures 1 through 9. In Figures 1, 2and 3 are shown the results from the root mean square prediction error (absolute mean prediction bias) analysis for CO, HC, and NOx respectively on a cycle basis. Figures 4, 5, and 6 show the results on a 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 MOBILE5a, but for NOx MEASURE performs equally well or slightly better than MOBILE5a. On a technology class basis, MEASURE is only marginally better than MOBILE5a for CO and HC, and results are mixed for NOx.

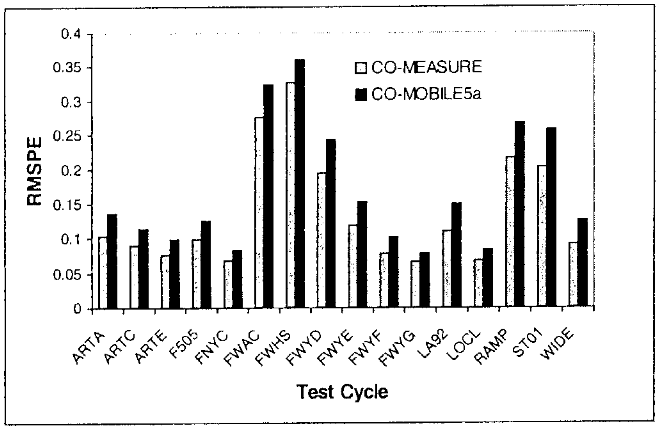

Figure1. Comparison of Root mean square prediction error of CO emission factors (or absolute mean prediction error) between MEASURE and Mobile5a.

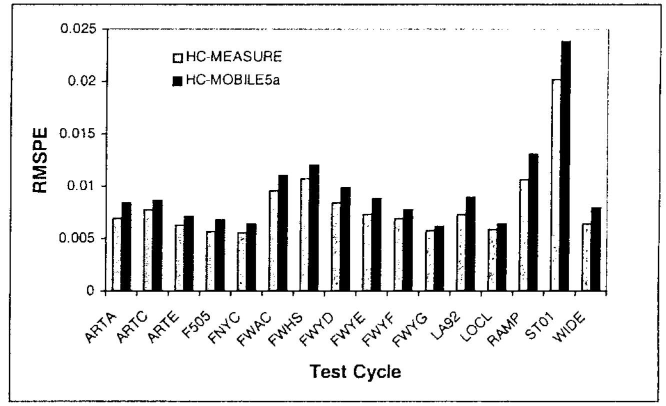

Figure2. Comparison of Root mean square prediction error of HC emission factors (or absolute mean prediction error) between MEASURE and Mobile5a.

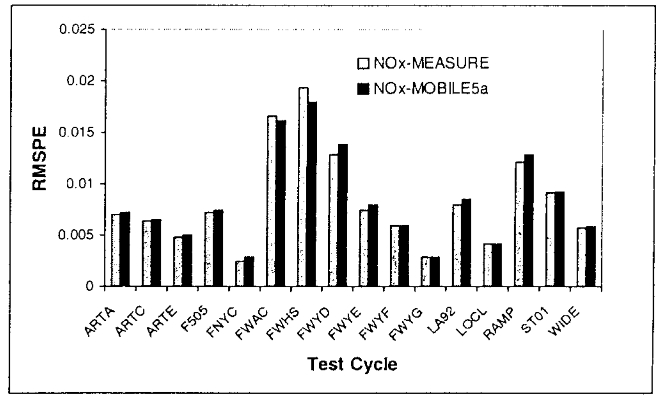

Figure 3. Comparison of Root mean square prediction error of NOx emission factors (or absolute mean prediction error) between MEASURE and Mobile5a.

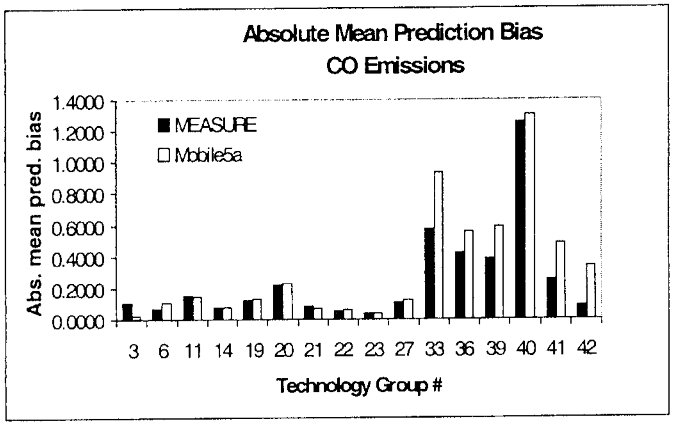

Figure 4. Differences in absolute mean prediction bias between MEASURE and Mobile 5a for predicted CO emission factors, by technology group.

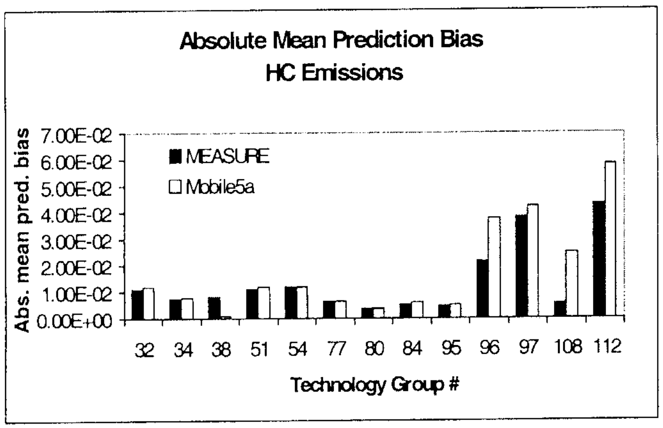

Figure 5. Differences in absolute mean prediction bias between MEASURE and Mobile5a for predicted HC emission factors, by technology group.

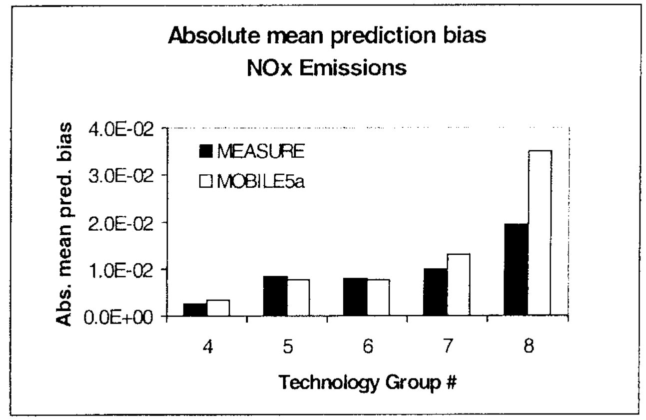

Figure 6. Differences in absolute mean prediction bias between MEASURE and Mobile 5a for predicted NOx emission factors

Figures 7, 8, and 9 show the results of the mean prediction bias (observed - predicted) on a technology group basis. When comparing mean prediction bias, it was found that MEASURE consistently over-predicts , while MOBILE5a consistently under-predicts, both on a cycle (Table 6) and technology class bases. Furthermore, as seen earlier, the same results indicate that across all cycles and technology classes the degree of over-prediction (as measured by the absolute mean prediction error) by MEASURE is lower than that of under-prediction by MOBILE5a, demonstrating once again by this measure of assessment that MEASURE performs better than MOBILE5a.

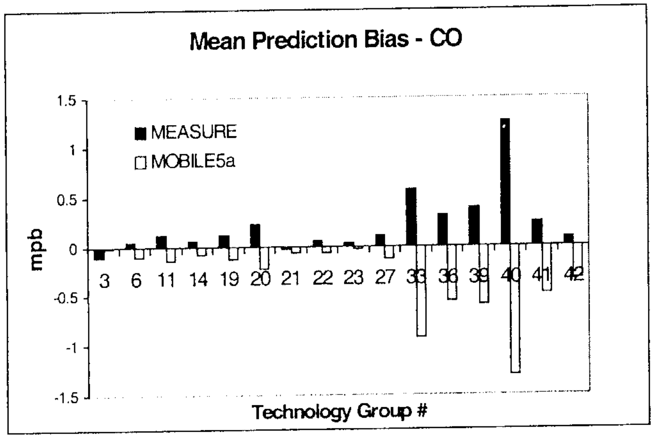

Figure 7, Differences in Mean Prediction Bias Between MEASURE and MOBILE5a Predictions on Technology Group Basis for Carbon Monoxide Emission Rates.

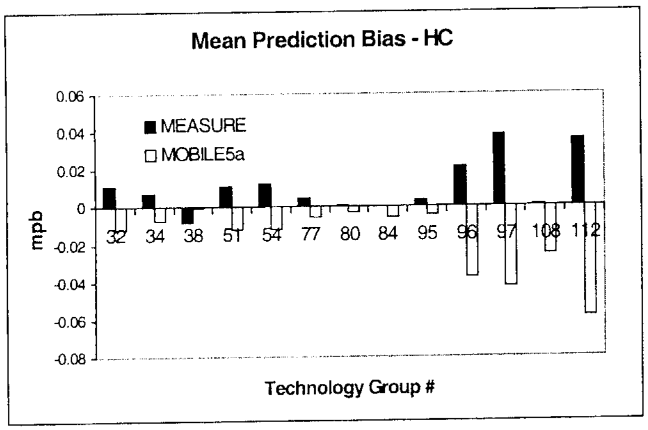

Figure 8, Differences in Mean Prediction Bias Between MEASURE and MOBILE5a Predictions on Technology Group Basis for Hydrocarbon Emission Rates.

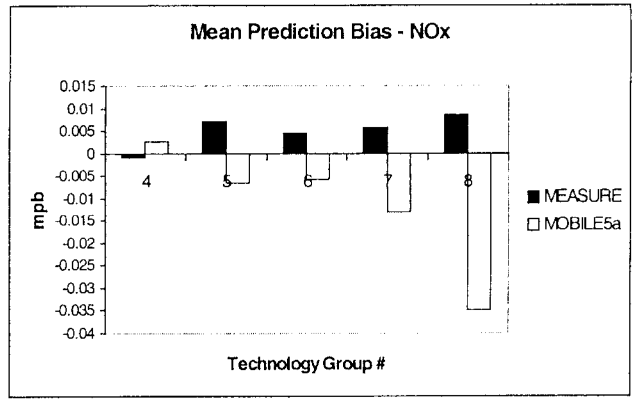

Figure9. Differences in Mean Prediction Bias Between MEASURE and MOBILE5a Predictions on Technology Group Basis for NOx Emission Rates

Table 6 Mean Prediction Bias: Observed vs. MEASURE, and Observed vs. MOBILE, By Cycle

Cycle	Observed Vs. Predicted- MEASURE (g/sec)			Observed Vs. Predicted - MOBILE 5a (g/sec)		
	CO	HC	NOx	CO	HC	NOx
ARTA	**0.0737**	**0.0042**	0.0062	-0.1344	-0.0078	0.0063
ARTC	**0.0539**	**0.0045**	0.0059	-0.1128	-0.0081	**0.0059**
ARTE	**0.0366**	**0.0030**	**0.0044**	-0.0956	-0.0068	0.0048
F505	**0.0637**	**0.0025**	**0.0061**	-0.1246	-0.0062	0.0063
FNYC	**0.0223**	**0.0021**	**0.0021**	-0.0815	-0.0063	0.0027
FWAC	**0.2646**	**0.0069**	**0.0015**	-0.3232	-0.0104	0.0110
FWHS	**0.3136**	**0.0082**	**-0.0007**	-0.3595	-0.0110	0.0105
FWYD	**0.1742**	**0.0054**	**0.0048**	-0.2424	-0.0093	0.0102
FWYE	**0.0891**	**0.0046**	**0.0061**	-0.1512	-0.0084	0.0067
FWYF	**0.0417**	**0.0036**	0.0054	-0.1001	-0.0072	**0.0052**
FWYG	**0.0170**	**0.0019**	0.0022	-0.0749	-0.0055	**0.0021**
LA92	**0.0866**	**0.0048**	**0.0073**	-0.1496	-0.0085	0.0079
LOCL	**0.0209**	**0.0022**	0.0037	-0.0786	-0.0059	**0.0035**
RAMP	**0.2005**	**0.0089**	**0.0111**	-0.2660	-0.0127	0.0120
ST01	**0.2030**	**0.0202**	**0.0091**	-0.2564	-0.0238	0.0092
WIDE	**0.0039**	**0.0039**	0.0052	-0.1243	-0.0075	**0.0052**
Weighted Average	*0.1079*	*0.0054*	*0.0050*	*-0.1671*	*-0.0091*	*0.0068*

Predictions biases in Table 6 appear to be relatively small, making their practical significance hard to gage. The difference in the prediction abilities of the two models is perhaps best illustrated by applying both models to the on-road fleet in the Atlanta Metro area, in a hypothetical scenario. There is approximately 100 million vehicle miles traveled per day in the Atlanta metropolitan region. Assuming that typical driving for this scenario is represented by the average of all driving cycles, which enables one to use average bias across cycles, the prediction bias per day of travel in the Atlanta region would be as shown in Table 7. That is, the MEASURE model over predicts CO emissions by 3795 tons per day, while Mobile 5a under predicts CO emissions by 5881 tons per day. Showing the numbers in tons per day of pollutant illustrates the importance of prediction biases—in aggregate they result in large errors in computing regional emission inventories.

Table7. Prediction Bias in Tons/Day for Average Daily Trip Making in Atlanta (MEASURE vs. Mobile5a)

Pollutant	MEASURE	Mobile5a
CO	3795	-5881
HC	191	-319
NOx	176	-241

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, that is, average speed works sufficiently well for NOx.

Some of the driving cycles used in the validation study were designed by EPA contractors to represent on-road 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 slightly over-predicted emissions for the validation set of vehicles, but this was not a huge concern, since it is preferable to err on the side of caution, and the bias is small. 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.

References

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

Fomunung, I, S. Washington, R. Guensler, and William Bachman. "Performance Evaluation of MEASURE Emission Factors, A Comparison to Mobile5a. Presented at the 79th Annual Meeting of TRB, National Research Council, Washington DC, January 2000.

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

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

Washington, S., Wolf, J., Guensler, R., 1997. A Binary Recursive Partitioning Method for Modeling Hot-Stabilized Emissions from Motor Vehicles. Transportation Research Record #1587, pp. 96-105. National Research Council Washington, D.C.

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.