ISSN 1047-3289 *J. Air & Waste Manage. Assoc.* **46**: 667-675

Motor Vehicle Emissions Variability

Gary A. Bishop and Donald H. Stedman
University of Denver, Denver, Colorado

Lowell Ashbaugh
University of California at Davis, Davis, California

ABSTRACT
Test-to-test variability has been observed by many current testing methods, including the Federal Test Procedure, the IM240 dynamometer test, the idle test common to many Inspection and Maintenance programs, and on-road remote sensing. The variability is attributable to the vehicle, not to the testing procedure. Because the vehicles are the dominant source of variability, the only way such vehicles can be reliably identified is through the use of multiple tests. The emissions variability increases with increasing average emissions, and it appears to be prevalent among the few newer technology vehicles with defective, but untampered, closed-loop emissions control systems (1981 and newer models). In one fleet the variable emitters constitute 2.2% to 4.8% of the vehicles and contribute 8.5% to 22% of the total carbon monoxide emissions. Scheduled I/M programs that fail to ensure repair of these vehicles allow a significant portion of vehicles with excess emissions to escape reduction measures.

INTRODUCTION
In the past 25 years the United States has spent billions of dollars to reduce emissions of new and in-use motor vehicles. For new vehicles this initiative has been an unqualified success; production vehicles now emit less than 4% of the hydrocarbons (HC) and carbon monoxide (CO) of their precontrol counterparts.[1] This dramatic reduction has not, however, had the anticipated result of reducing urban ozone levels in many U.S. cities. A body of evidence is growing that implicates the performance of the in-use vehicle fleet as the reason these air quality goals are not being met.[2-6]

> **IMPLICATIONS**
> Variable exhaust emissions are observed from malfunctioning motor vehicles by all testing methods, including the Federal Test Procedure. These vehicles are a significant emissions source; the variable nature of their emissions will allow them to pass scheduled Inspection and Maintenance testing programs, thereby avoiding corrective repairs.

In areas of the country that violate the federal air quality standards, the in-use vehicle fleet is required to submit to some type of emissions inspection program. The goal of these programs is to maintain adequate emissions performance as the vehicles age. These inspection and maintenance (I/M) programs—which occur annually, biennially, or on change of ownership—range from anti-tampering programs performed at decentralized garages to annual tailpipe and emissions component integrity inspections performed by a central contractor. Owners whose vehicles fail these inspections are required to comply with the local regulations before they can register or transfer titles of their vehicles. U.S. Environmental Protection Agency computer models assume that compliance for failed vehicles includes some type of emissions repair with long lasting benefits.[7] The data suggest otherwise.[8]

Recent studies suggest that many vehicles are escaping the repair component of emission testing programs and returning to the highways without needed repairs.[9,10] Lawson examined roadside data collected in California, and showed that fleet tampering rates for emissions control equipment are similar between areas that have biennial inspections and those that do not have the inspections.[11] He showed that overall I/M failure rates for vehicles stopped on California highways are 1.8 times higher than the failure rate reported for the biennial inspection. He also showed that their roadside emissions were significantly higher than those reported for their smog check inspection. This "clean for a day" phenomenon is in stark contrast to controlled repair studies that show repairs generally last in excess of three years.[12] These data suggest that vehicles that should have received lasting repairs through the I/M program are somehow circumventing the process. This severely limits the benefits derived from scheduled I/M programs. Several illegal actions can be taken by owners who choose to avoid repairs; however, test-to-test variability is an entirely legal way that some vehicles avoid repairs.

With the successful demonstration of on-road remote sensing devices (RSD) for vehicle exhaust, one can now routinely perform multiple measurements on in-use vehicles. Data collected with these instruments have drawn criticism for displaying highly variable emissions on measurements made a few seconds apart.[13,14] Figure 1 shows typical data

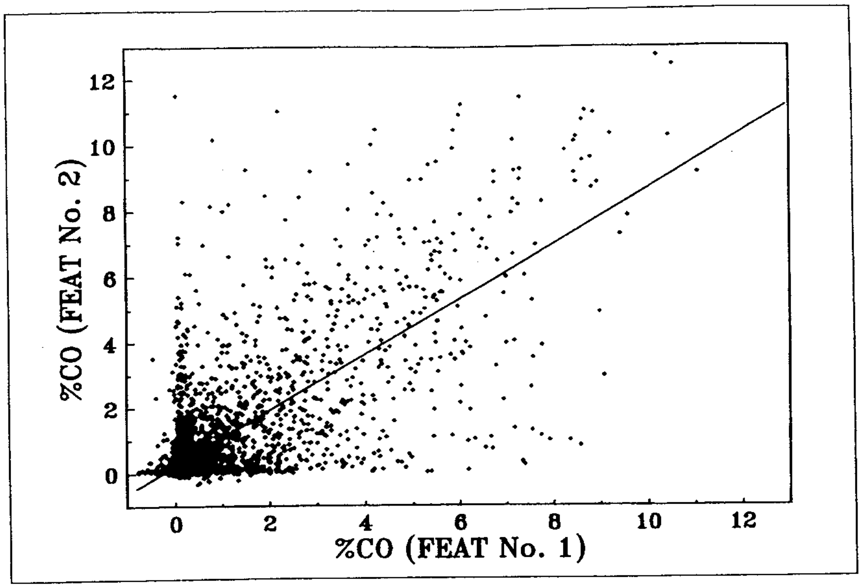

Figure 1. Data collected from 4,122 vehicles on Rosemead Boulevard in El Monte, CA using two FEAT units approximately 100 feet apart. The equation of the regression line is FEAT(2) = 0.23 + 0.85*FEAT(1), with r^2 = 0.54.[15]

collected in California from two sensors located approximately 100 feet apart.[15] Initially, the concern focused on the validity of the measurements; i.e., the lack of correlation was thought to result from the inability of a remote sensor to accurately measure the instantaneous exhaust emissions. However, collocated remote sensing measurements and double blind comparisons to vehicles with onboard emissions monitors[16-18] have resulted in CO and HC correlations with r^2 typically greater than 0.9.

It seems reasonable that a one-second sample should exhibit higher test-to-test variability than lengthier tests. This will certainly be true if the sampling method is a major source of the variability. However, results from California and other studies have suggested that, for some vehicles, emission variability is intrinsic to the vehicle. If this is correct, then variability similar to Figure 1 should also be exhibited when other emission tests are repeated on representative fleets.

APPROACH

Emissions data were collected from a number of different studies in which the same emissions test was conducted more than once on the same vehicle. The tests examined include the Federal Test Procedure (FTP), the IM240, no-load idle measurements, and on-road remote sensing. Test data were analyzed to reveal the emissions variability as a function of emission magnitude. It is not surprising that the low emitting vehicles exhibited low variability. The only way a vehicle can show low emissions by our methods is to be consistently low. On the other hand, we found that some high emitters show high variability no matter what testing procedure is used. This finding is surprising, as we expected the longer tests to show more consistent measurements from one test to another.

Since the early 1970s regulatory agencies have required each preproduction vehicle/drive train combination sold in the United States to have exhaust emissions certified to various limits using the Federal Test Procedure.[19-22] Each test takes at least 12 hours to complete and costs more than $1,000. Precision of the results for a given vehicle has been reported to be ±20% and is controlled mainly by the reproducibility of the automobile's emission system, not by the test system or gas analysis protocols.[23] Vehicles are certified by remaining below emission standards for two consecutive tests. The number of failed tests does not count against the vehicle in the certification process. The results of the FTP have been used as the basis for computer models of on-road emissions, even though the test was not designed for that purpose.

The expense and time requirements of the Federal Test Procedure preclude its use in vehicle inspection programs. For use in these programs, federal and state regulatory agencies have designed shorter, less expensive tests that can be used on the millions of in-use vehicles on the road today. One such test, the IM240, consists of a 240-second cycle on a dynamometer that includes accelerations, cruise, and decelerations. Many states use a no-load idle test in their vehicle testing programs. Some of these tests include a low idle (~1,000 rpm) and a high idle (2,500 rpm).

The ability of the shorter tests to faithfully reproduce FTP results has been used as a measure of their success. In fact, the FTP measurements are held in such high regard that:

> Correlation with the FTP is critical for any test procedure that might be used to trigger vehicle maintenance requirements. The FTP is known to be a "representative" driving cycle in terms of average speed, stops per mile, major speed deviations per mile, and minor speed deviation pattern.[14]

The acceptance of the FTP as a representative test has led to the widespread belief that FTP measurements are invariant. This conviction has been reinforced by the fact that FTP measurements are rarely duplicated on the same vehicle, especially on high emitting vehicles. Because of the long driving cycle and careful preconditioning, variability was thought to be eliminated or reduced to the point of being irrelevant.

In 1992 a consortium of automobile manufacturers and oil companies undertook the Air Quality Improvement Research Program (AQIRP) to examine the effects of many of

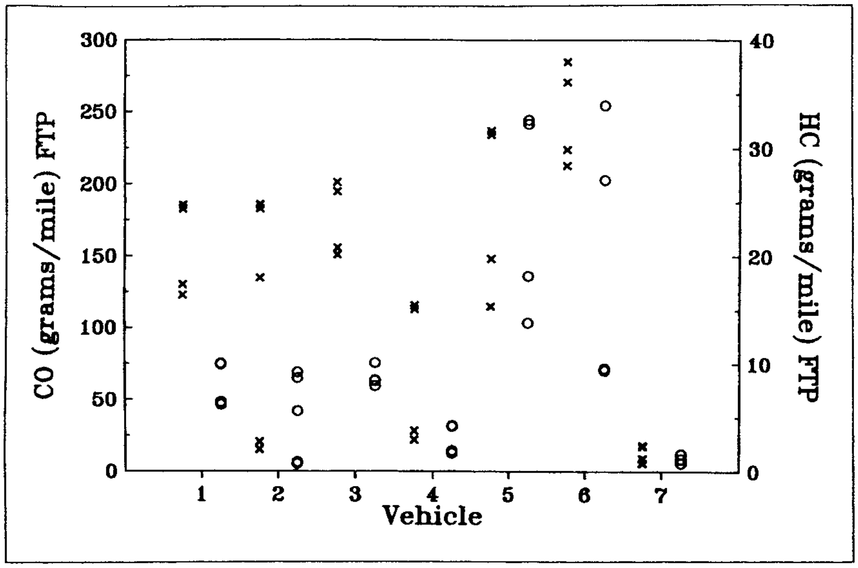

Figure 2. FTP data for CO and HC emissions from seven 1986 and newer model year high emitters. Four (or five) separate tests on the same fuel (studies base fuel A) are plotted for each vehicle for CO (x) and HC (o).[24]

the proposed fuel modifications outlined in the Clean Air Act Amendments of 1990 on exhaust emissions from late model cars.[24,25] Vehicles were recruited and segregated into two categories: normal emitters and high emitters. The high emitter study recruited nine vehicles, defined by AQIRP as 1986 and later model year vehicles with untampered emission control systems and with initial IM240 failure emissions of CO greater than 15 grams per mile and/or HC greater than 1.1 grams per mile. Upon delivery to the research laboratory, confirmatory FTP testing eliminated two of the nine vehicles from the study for failing to meet this definition. The remaining vehicles were tested numerous times on a variety of fuels by the FTP. We used only the tests performed on the base fuel (denoted as fuel A in the report) from these seven vehicles in our analysis.

The California Air Resources Board (CARB) measured emissions from 554 vehicles at its El Monte laboratory[26] using a remote sensor and a dynamometer test. Each vehicle received three remote sensor measurements and one IM240 measurement (CARB converted the IM240 results to FTP equivalent using a linear relationship). We compared data from the first two valid RSD measurements to the "FTP" emissions in grams per mile.

The state of Delaware measured emissions two times from 20 vehicles on the IM240 test and once on the FTP in its Vehicle Retirement Program.[27] We also obtained a similar data set on 213 vehicles recruited by the U.S. EPA.[28] We compared CO and HC emissions from two separate IM240 dynamometer tests performed on each vehicle.

The Southwest Research Institute collected idle test data for the U.S. EPA in 1987.[29] The tests included idle and 2,500 rpm emissions for 25 fully warmed-up vehicles measured weekly upon arrival at work over a fifteen-week period. A box-and-whisker plot is used to compare the emissions variability found in the CO and HC data.

Finally, we examined our own measurements of CO and HC from 3,624 vehicles collected on Rosemead Boulevard in El Monte, CA for which we could find three or more remote sensing measurements. We also analyzed roadside pullover inspections from 111 vehicles collected at the same time.[15] We calculated the average CO and HC exhaust concentrations and the variance for each vehicle, divided the data set into deciles by average emissions, and computed the average square root of the sample variance within each decile.

RESULTS

Figure 2 shows the AQIRP measurements of carbon monoxide and hydrocarbon emissions from four FTP tests (archived data includes one additional test for vehicle #2) for the seven test vehicles. The absolute test-to-test variability for these repeat measurements is quite high, with the worst cases varying by more than an order of magnitude. These data raise the question of which FTP test represents the "true" emissions of these vehicles—the highest, the lowest, or an average of all of them? This is extremely important, as the average HC and CO emissions from these vehicles are 50 to 100 times those of the normal emitting control vehicles. The emissions variability of these vehicles is much larger in magnitude (and in percentage) than the relatively subtle effects of fuel chemistry variation which was the focus of the AQIRP study.[24]

The FTP cycle was designed to ensure that each measurement accurately reflects the emissions of that vehicle at the time of the test. However, as noted by the AQIRP researchers, much of the observed variability "was a direct consequence of unstable vehicle operation that resulted in poor control of air:fuel equivalence ratio."[24] It is apparent that the vehicle, not the test, is the dominant source of the large observed test-to-test variability.

Figures 3A and 3B show the data from the 554 vehicle fleet measured by the California Air Resources Board at its El Monte laboratory.[26] We have plotted the first two valid RSD measurements versus the log of the "FTP" emissions in grams per mile for CO (a total of 542 vehicles) and for HC (a total of 538 vehicles). The average FTP emissions for the entire fleet was 26 grams per mile for CO and 2.5 grams per mile for HC. The vertical axis shows the two separate remote sensor exhaust measurements, which were recorded on a flat and level roadway at a constant speed of 20 mph. The majority of measurements show low FTP emissions and little RSD test-to-test variability. As the FTP emissions increase,

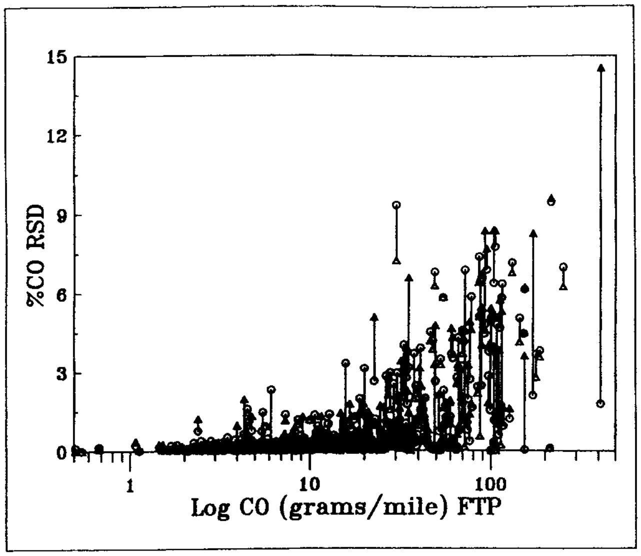

Figure 3A. California Air Resources data for 542 vehicles measured twice for %CO by an RSD (o indicates first measurement and Δ represents the second measurement) at constant speed and load plotted versus LOG CO emissions in grams per mile.[26]

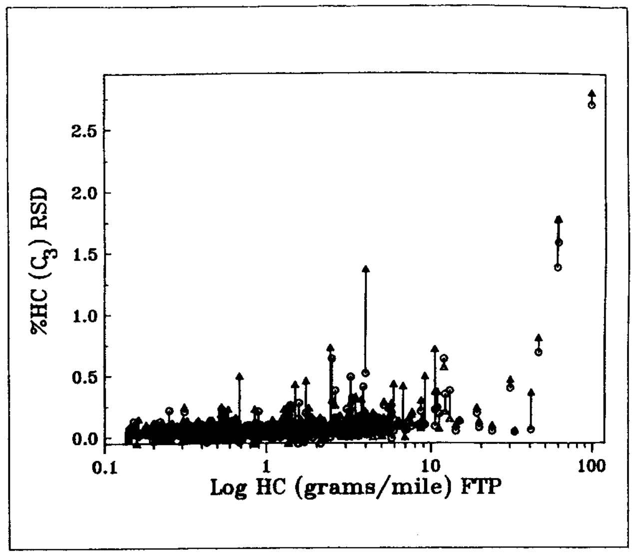

Figure 3B. California Air Resources data for 538 vehicles measured twice for %HC (propane) by an RSD (o indicates first measurement and Δ represents the second measurement) at constant speed and load plotted versus LOG HC emissions in grams per mile.[26]

the variability of the remote sensing measurements also increases; the onset occurs around 20 grams per mile for CO. Hydrocarbon measurements show less variability than CO, in part due to the smaller number of high emitting HC vehicles. The variability observed by the remote sensor in this study is consistent with the observed FTP variability of the high emitters illustrated in Figure 2.

Figure 4A shows combined CO emissions data from the 20 State of Delaware Vehicle Retirement Program vehicles and the 213 vehicles recruited by the U.S. EPA. The figure compares CO emissions from two separate IM240 dynamometer tests performed on each vehicle. For the EPA data set, (Δ) represents tests performed at its IM240 lane in Hammond, IN and (o) represents tests performed by an EPA contractor in its emissions laboratory. As in Figure 3A, the vehicles are listed along the x-axis by increasing CO gram/mile FTP emissions. For the 233 vehicles, the lowest emitter is 0.62 grams per mile CO, and the highest is 271.82 grams per mile CO. The average for the entire fleet is 28.4 grams per mile CO. The onset of variability occurs at an emissions level of approximately 20 grams per mile CO, similar to what was found by the RSD.

Figure 4B shows the hydrocarbon emissions data for the same vehicles as for CO in Figure 4A. The main difference is that there are fewer gross polluting vehicles for hydrocarbon than for carbon monoxide; however, a large test-to-test

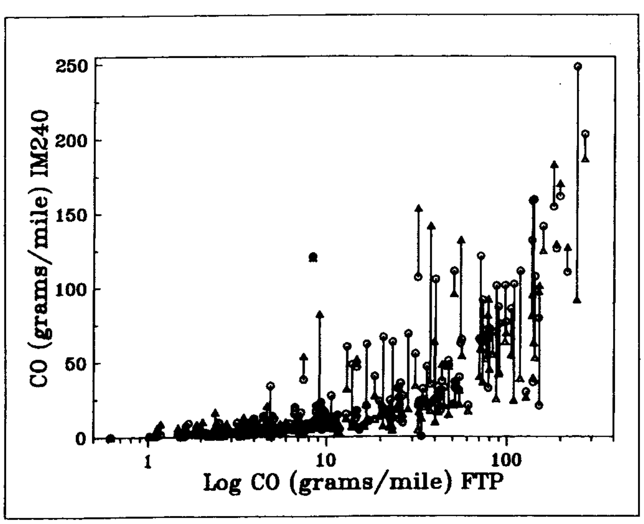

Figure 4A. Combined data from 233 vehicles from the U.S. EPA and the State of Delaware's Vehicle Retirement program where two IM240 measurements for CO are made for each vehicle and plotted versus the log of its CO emissions as determined by FTP.[27,28]

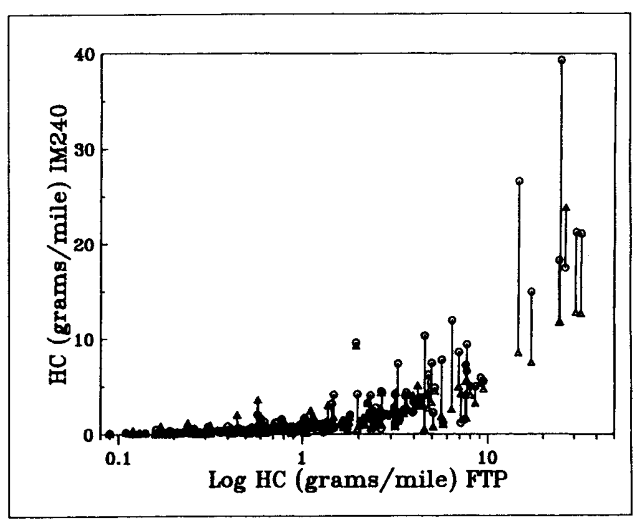

Figure 4B. Combined data from 233 vehicles from the U.S. EPA and the State of Delaware's Vehicle Retirement program where two IM240 measurements for HC are made for each vehicle and plotted versus the log of its HC emissions as determined by FTP.[27,28]

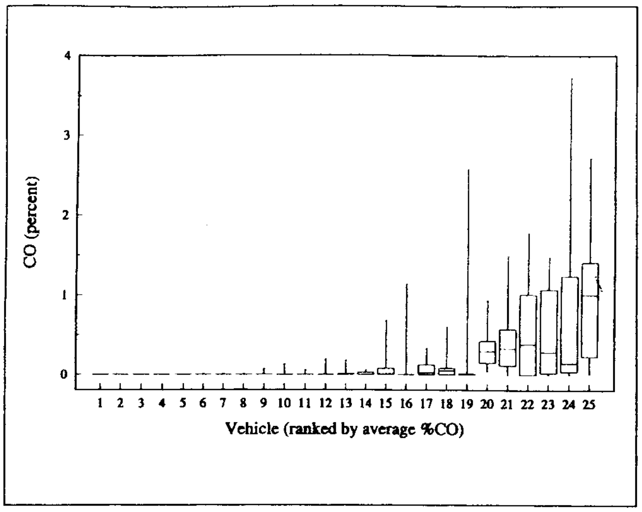

Figure 5A. A box-and-whisker plot for 25 vehicles which underwent weekly idle/2,500 rpm idle testing upon arrival at work for a 15-week period. The vehicles are rank ordered by their combined mean %CO idle/2500 rpm measurements along the x-axis and the box is drawn to represent emission values for the 25th, 50th and 75th percentiles with the whiskers covering the extent of the data.[29]

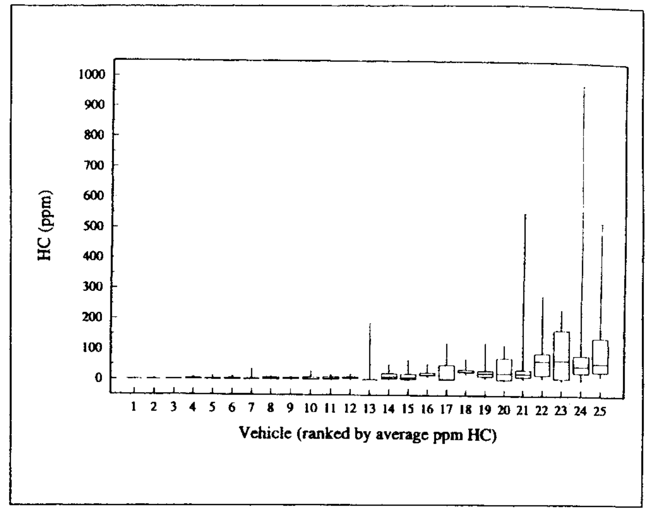

Figure 5B. A box-and-whisker plot for 25 vehicles which underwent weekly idle/2,500 rpm idle testing upon arrival at work for a 15-week period. The vehicles are rank ordered by their combined mean ppm HC (as hexane) idle/2500 rpm measurements along the x-axis and the box is drawn to represent emission values for the 25th, 50th and 75th percentiles with the whiskers covering the extent of the data.[29]

variability is still observed among the highest emitting vehicles. The FTP HC emissions range from a low of 0.09 grams per mile to a high of 32.6 grams per mile; the average for this data set is 2.24 grams per mile.

Figures 5A and 5B display the weekly idle test data collected by the Southwest Research on 25 fully warmed-up vehicles measured upon arrival at work over a fifteen week period. A box-and-whisker plot is used to show the observed measurement variability with the dimensions of the box defined by the 25th, 50th, and 75th percentiles and the whiskers covering the extent of the data. The vehicles are rank ordered by increasing average idle/2500 rpm exhaust measurements, which for CO ranges from 0% to 0.9%, and for HC between 0.2 ppm to 107.1 ppm (hexane). The graphs show levels of emission variability which are similar to the previous figures.

Figures 6A and 6B show the emissions variability characteristics of CO and HC, respectively, from the 3,624 vehicles we measured by remote sensing on Rosemead Boulevard in El Monte, CA in June of 1991. The average CO and HC exhaust concentrations and the square root of the sample variance were calculated for each vehicle; the data are displayed by deciles of average emissions. The average for each decile is plotted as a horizontal line, and the vertical bar represents

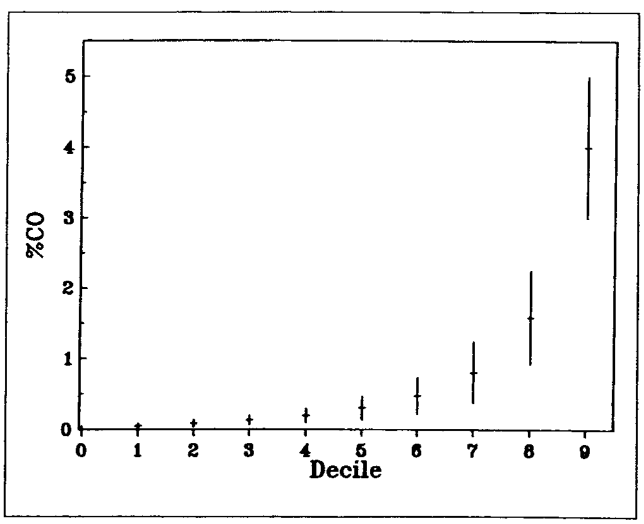

Figure 6A. Carbon monoxide emissions by decile for 3,624 vehicles measured three or more times on Rosemead Boulevard in El Monte, CA.[15] The average %CO emissions for each decile are plotted as the horizontal bar with the vertical line being equal in length to the average square root of the sample variance.

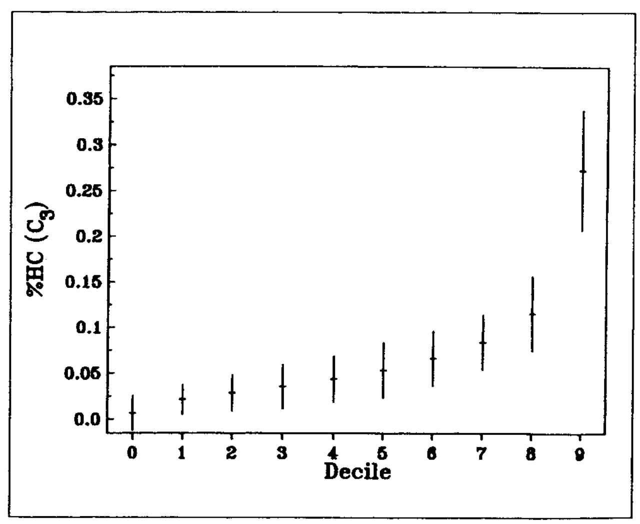

Figure 6B. Hydrocarbon emissions by decile for 3,624 vehicles measured three or more times on Rosemead Boulevard in El Monte, CA.[15] The average %HC (as propane) emissions for each decile are plotted as the horizontal bar with the vertical line being equal in length to the average square root of the sample variance.

the average square root of the sample variance for each decile. Both the CO and HC plots show that the square root of the sample variance increases as average emissions increase; the last decile has the highest variability. Measured by grams per gallon of fuel burned, the last decile (10% of the vehicles) contributed 53% of the CO emissions and 37% of the hydrocarbon emissions. In contrast, the eight lowest emitting deciles (80% of the vehicles) accounted for 47% of the fleet HC emissions and only 27% of the CO emissions. On average these 3,624 vehicles emitted exhaust concentrations of 0.77% CO and 0.073% HC (as propane); the mean model year was 1985, and these results were statistically similar to all of the 42,546 measurements collected on Rosemead Boulevard.

Table 1 summarizes the above data along with data from the Colorado Department of Health comparing FTP measurements to the other dynamometer short tests that are reported to favorably correlate with the FTP.[32,33] The data sets are ordered according to increasing average FTP emissions for each pollutant species. The coefficient of determination decreases as the average FTP emissions increase, indicating the higher variability that occurs in the higher emitters. The relationship between IM240 data collected at the EPA test lane and the EPA contractor's laboratory (reference 3) is similar to that displayed for remote sensing in Figure 1. For the combined data sets (314 vehicles), half of the FTP CO and HC emissions are contributed by 10.5% and 10.8% of the vehicles, respectively, which is also similar to the RSD data.

DISCUSSION

There are at least five aspects of automotive exhaust emissions variability that are important. First, test-to-test emissions variability has similar characteristics for all current test methods. This includes Federal Test Procedure (FTP) testing, remote sensing measurements, the related dynamometer short tests, and idle testing. This implies that the vehicle, not the testing method, is the dominant source of the variability. Second, since the vehicle is the dominant source of the variability a "better" test cannot be devised to eliminate the variability. Furthermore, only through the use of multiple tests can these vehicles be identified. Third, vehicle emissions variability increases with increasing overall emissions. Restated, low-emitting vehicles on average exhibit little test-to-test variability, while high-emitting vehicles can have very large changes in emissions from one measurement to another (CO changes on the order of 150 grams per mile and HC changes on the order of 30 grams per mile were observed in FTP testing). Fourth, we will discuss some similarities found among these vehicles and show that their emissions contribution is significant. Finally, the existence of high-emitting vehicles with variable emissions has profound implications for current and future I/M programs if they are to be successful in reducing on-road emissions.

It is not clear from these data why all test methods show similar variability characteristics. We know that variably emitting vehicles can have different time constants; i.e., high frequency "flippers" have high second-to-second emission variations, and low frequency "flippers" change more slowly, perhaps daily or weekly. Support for the existence of low frequency "flippers" can be drawn from the AQIRP data (see Figure 2) where several vehicles showed markedly different FTP emissions behavior on different days. Remote sensing data sets have consistently revealed the existence of high frequency variable emitters.[30,31] It is reasonable to expect that a high frequency "flipper" will have a greater impact on RSD measurements than on a dynamometer measurement; one possible explanation for the observed consistency is that high and low frequency "flippers" occur with similar frequency in the fleet and their emissions variability is distributed similarly. For example, assume we have a four-vehicle fleet whose tailpipe CO emissions are (1) 0%, (2) 8%, (3) high frequency (second-by-second) flipper between 0% and 8%, and (4) low frequency (daily) flipper between 0% and 8%. This fleet would be appear to contain only one variably emitting vehicle, number 3, for a single day multi-pass RSD series of measurements. Conversely, on any given day, an IM240 measurement would report no variable emitters; however, for measurements made on different days (like the data analyzed here) it becomes very likely that vehicle number 4 will be observed as a variable emitter. Thus each method could observe a similar percentage of variably emitting vehicles, with a similar percentage of emissions variability (25%), although for different vehicles.

Our analysis shows that vehicles with higher than average emissions have a greater probability of higher emissions variability; thus, all tests will have equal difficulty consistently detecting variable high-emitting vehicles. The

Table 1. Correlations of dynamometer short tests with FTP results as a function of fleet average FTP emissions.

Source	Species	Average FTP Emissions (g/mile)	Short Test	Number of Vehicles	r^2
CDH[1]	HC	2.0	CDH226	81	0.86
EPA[2]	HC	2.2	IM240	213	0.91
EPA[3]	HC	2.2	IM240	213	0.84
DVRP[4]	HC	7.2	IM240	20	0.75
CDH	CO	27.3	CDH226	81	0.66
EPA[2]	CO	28.4	IM240	213	0.73
EPA[3]	CO	28.4	IM240	213	0.62
DVRP	CO	80.5	IM240	20	0.32
EPA[2]	NO_x	1.3	IM240	213	0.80
EPA[3]	NO_x	1.3	IM240	213	0.73
CDH	NO_x	1.8	CDH226	81	0.73
DVRP	NO_x	2.2	IM240	20	0.32

[1]Colorado Department of Health data.[34]
[2]EPA Data, Laboratory performed both tests.[28]
[3]EPA Data, Short test performed by IM240 Lane at Hammond, IN.[28]
[4]Delaware Vehicle Retirement Program, non-waivered vehicles.[27]

Table 2. Frequency of on-road emissions variability as a function of vehicles underhood emissions equipment status.

Underhood Inspection Results	Number of Vehicles	Vehicles with Variable Emissions1	Percentage with Variable Emissions 1	Mean Model Year of Variable Emitters
Passed	51	22	43%	1982.5
Defective	21	9	43%	1981.2
Tampered	39	7	18%	1980.7
Totals	111	38	34%	1981.9

[1]A vehicle is defined as having variable emissions if one on-road RSD measurement exceeded the pullover cutpoints of 4% CO and 0.3% HC (propane), and at least one additional on-road measurement was less than or equal to half of these cutpoints

degree to which vehicles will be detected is a function of the emissions cutpoints used. The only option remaining to identify these high-emitting vehicles is to make use of the variability of their emissions and of the consistency of low emitters. Multiple testing of vehicles is unlikely to incorrectly identify low emitters as high emitters, but will detect high emitters whether they are consistent or variable.

We used the Rosemead Boulevard data collected in 1991 to find additional commonalities among vehicles with variable tailpipe emissions.[15] We defined a variable emitter as a vehicle with at least one RSD measurement above the high emitter pullover cutpoints (4% CO and 0.3% HC [as propane]) and one measurement less than or equal to half of the cutpoints. Out of 334 vehicles given a roadside Smog Check, we identified 111 vehicles that had been measured at least twice by the RSD. The vehicles on-road emissions behavior as a function of their underhood inspection results is shown in Table 2. Vehicles which showed evidence of deliberate tampering were older on average and had a lower incidence of variable on-road emissions than either of the other two categories.

In the AQIRP study, all seven high-emitting vehicles studied were modern closed-loop computer controlled vehicles that had not been tampered with. All of the vehicles were diagnosed to have at least one malfunctioning or broken control component or subsystem. Six of the seven vehicles had fault codes indicating problems with the oxygen sensor. It is not surprising that improper operation of the oxygen sensor feedback control could contribute to variable vehicle operation.

At the University of Denver we have recently completed a small high emitter remote sensing identification and repair program in which faulty oxygen sensors were also found to be responsible for high emissions variability.[35] Figure 7A shows second-by-second CO data for a 1984 Nissan Sentra driven over a five-mile street/highway course using an inflight four-gas analyzer (model 3400, OTC Division, SPX Corp, Owatonna, MN) before any repairs were attempted. The figure illustrates the high frequency variability of this vehicle's CO emissions. All of the CO values below 1% occurred when the vehicle was at idle. This observation is consistent with the fact that this vehicle had recently passed its Colorado I/M idle test. Repairs included a new oxygen sensor and catalytic convertor (removal and visual inspection of the old catalyst showed ~30% loss of substrate and substantial melting and plugging of the remaining matrix). Figure 7B is a similarly collected data trace from a 1988 Mercedes with a malfunctioning oxygen sensor. At times the system functioned correctly; for example, very low emissions were

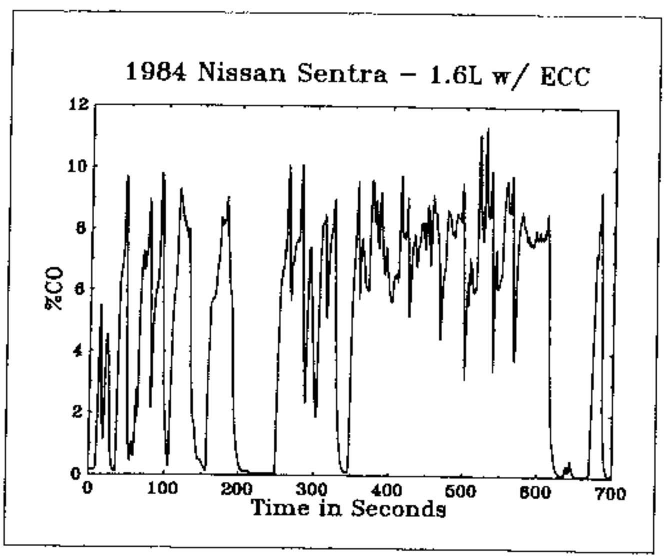

Figure 7A. Second-by-second %CO data collected with an in-flight four-gas analyzer from a 1984 Nissan Sentra with an electronically controlled carburetor. The trace covers a five-mile road course of stop-and-go/freeway driving in and around the University of Denver. Periods of low CO emissions occur at idle for this vehicle.[35]

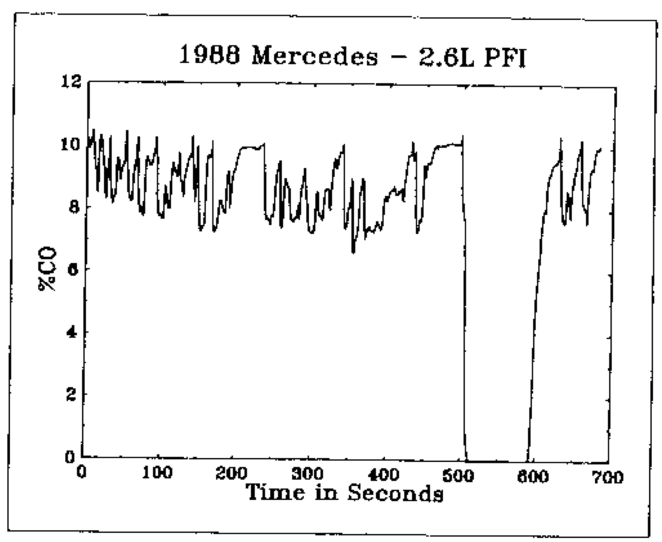

Figure 7B. Second-by-second %CO data collected with an in-flight four-gas analyzer from a 1988 Mercedes Benz 190E with port fuel injection. The trace covers a five-mile road course of stop-and- go/freeway driving in and around the University of Denver. Low emissions are observed during only one interval when presumably the vehicles oxygen sensor was functioning properly.[35]

observed between 500 and 600 seconds. The only repair necessary for this vehicle was replacement of the oxygen sensor. This type of failure could produce low frequency emissions variability depending on the frequency and duration with which the oxygen sensor is able to function properly. The "check engine" light that is supposed to indicate an oxygen sensor malfunction was not illuminated on either vehicle.

These three studies suggest that while tampered vehicles may experience high on-road emissions, the tampering may decrease variability. Non-tampered high emitting vehicles appear to be more likely to show variable emissions, especially when an oxygen sensor problem is involved.

To estimate the possible overall contribution to fleet emissions produced by variable emitting vehicles the remote sensor CO data from Figure 1 were analyzed using three different definitions for a variably emitting vehicle. To be classified as variable emitters, all vehicles were required to exceed the RSD pass/fail criteria of 4% CO at least once and to have a second reading which was I) 50% below the pass/fail mark, II) 25% below, or III) simply less than the pass/fail criteria. Using these definitions, the percent of the fleet with variable emissions, as measured by the two remote sensors, is 2.2% for definition I, 3.4% for definition II, and 4.8% for definition III. The percentage of the total exhaust mass emissions (on a grams per gallon basis) is 8.6% for definition I, 14.3% for definition II, and 22.2% for definition III.

These observations all have important implications for Inspection and Maintenance programs. Probably the most important is that a single high emissions reading on a fully warmed-up vehicle, not observed during periods of hard accelerations, likely indicates a vehicle which is in need of some type of repair. On the other hand, a single low reading, regardless of test type, is not sufficient evidence of proper repair. Vehicle #2 in Figure 2 operates at a very low emissions level about half of the time. This low emissions level would satisfy most of the new and more stringent I/M program requirements, yet the vehicle needs repair. Also vehicles found in the last decile of Figure 6 are often observed below the high emitter cutpoints. This points out that infrequent testing, and certification after a single measurement below the cutpoint, provides a way for variable high emitting vehicles to escape I/M programs without repair. The current movement toward less frequent testing[27,36,37] will allow these vehicles to remain on the road longer without repair, and may increase the fleet average emissions as a result.

CONCLUSIONS

We have documented variable test-to-test vehicle exhaust emissions, regardless of the type of emissions test. This supports the notion that the vehicle, and not the test method, is the dominant source of exhaust emissions variability. The variability increases with increasing average emissions, and because of the erratic nature of some vehicles it cannot be eliminated with longer test averaging times. However, the possibility exists that different testing methods are observing different vehicles exhibiting statistically similar emissions variability.

Test-to-test emissions variability is a problem for current infrequently scheduled vehicle Inspection and Maintenance programs. Dynamometer driving cycles like the Federal Test Procedure and IM240 were developed to average emissions over a long enough period of time (and over enough operating conditions) to provide results with low variability. Upon careful examination, however, the evidence shows that malfunctioning vehicles that are among the highest emitters often show very inconsistent FTP or IM240 emissions.

We have shown that variable emitters can be identified only through more frequent testing, not by using a more sophisticated test. In fact, if a variably emitting vehicle has fewer chances to be detected as a high emitter, it has a higher probability of escaping detection and avoiding repair.

Our analysis suggests that vehicles with highly variable on-road tailpipe emissions are likely to be modern computer-controlled vehicles that have malfunctioning emission control systems (especially oxygen sensor problems) that have not been tampered with. They are likely to be overall high emitters that contribute significantly to excess on-road emissions. They also exhibit periods of operation that could enable them to pass an I/M inspection. Thus, for any type of I/M testing, as long as one passing test determines compliance, irrespective of previous failures, the potential exists for a significant fraction of high emitters to escape without needed repairs. This knowledge should help I/M program administrators focus their attention on ways to identify these vehicles and improve their programs.

REFERENCES

1. Calvert, J.G.; Heywood, J.B.; Sawyer, R.F.; Seinfeld, J.H. "Achieving acceptable air quality: Some reflections on controlling vehicle emissions," *Science* 1993, *261*, 37.
2. National Research Council Committee on Tropospheric Ozone Formation and Measurement. *Rethinking the Ozone Problem in Urban and Regional Air Pollution*, National Academy of Sciences, National Academy Press, Washington, D.C., 1991.
3. Ingalls, M.N.; Smith, L.R.; Kirksey, R.E. "Measurement of on-road vehicle emission factors in the California South Coast Air Basin - Volume I: Regulated emissions." Final Report No. SWRI-1604 from Southwest Research Institute to the Coordinating Research Council, Atlanta, GA, June 1989.
4. Pierson, W.R.; Gertler, A.W.; Bradow, R.L."Comparison of the SCAQS tunnel study with other on-road vehicle emission data," *J. Air & Waste Manage. Assoc.* 1990, *40*, 1495.
5. Fujita, E.M.; Croes, B.E.; Bennett, C.L.; Lawson, D.R.; Lurmann, F.W.; Main, H.H. "Comparison of emission inventory and ambient concentration ratios of CO, NMOG, and NO_x in California's south coast air basin," *J. Air & Waste Manage. Assoc.* 1992, *42*, 264.
6. Fujita, E.M.; Watson, J.G.; Chow, J.C.; Lu, Z. "Validation of the chemical mass balance receptor model applied to hydrocarbon source apportionment in the southern California air quality study," *J. Air & Waste Manage. Assoc.* 1994, *28*, 1633.
7. Glover, E.L.; Brzezinski, D.J. "MOBILE4, Exhaust emission factors and inspection/maintenance benefits for passenger cars," U.S. Environmental Protection Agency, Ann Arbor, 1989; EPA-AA-TSS-I/M-89-3.
8. McCargar, J.A.; Snapp, L.M. "Report on the EPA/Manufacturer Cooperative I/M Testing Program," U.S. Environmental Protection Agency, Ann Arbor, 1992; EPA/AA/EPSD-I/M-92-01.

9. Ashbaugh, L.L.; Lawson, D.R. "A comparison of emissions from mobile sources using random roadside surveys conducted in 1985, 1987, and 1989," Air & Waste Management Association, Vancouver, B.C., June 1991; Paper No. 91-180.58.
10. Beaton, S.P.; Bishop, G.A.; Zhang, Y.; Ashbaugh, L.L.; Lawson, D.R.; Stedman, D.H. "On-road vehicle emissions: Regulations, costs, and benefits," *Science* **1995**, *268*, 991.
11. Lawson, D.R. "'Passing the test' - human behavior and California's smog check program," *J. Air &Waste Manage. Assoc.* **1993**, *43*, 1567.
12. Brown, N.D.; Glover, E.L.; Clemmens, W.B. "Effect of engine condition on FTP emissions and in-use repairability," Society of Automotive Engineers; Paper No. 940302; 1992.
13. Austin, T.C.; DiGenova, F.J.; Carlson, T.R. "Analysis of the effectiveness and cost-effectiveness of remote sensing devices," report to the U.S. Environmental Protection Agency, Sierra Research, Inc., 1994.
14. Austin, T.C.; Carlson, T.R.; Gianolini, K.A. "An evaluation of 'remote sensing' for the measurement of vehicle emissions," report to the California Air Resources Board and the California I/M Review Committee, Sacramento, August 28, 1990; Paper No. SR90-08-02.
15. Stedman, D.H.; Bishop, G.A.; Beaton, S.P.; Peterson, J.E.; Guenther, P.L.; McVey, I.F.; Zhang, Y. "On-road remote sensing of CO and HC emissions in California," Final Report, Contract #A032-093, California Air Resources Board Research Division, February, 1994.
16. Ashbaugh, L.L.; Lawson, D.R.; Bishop, G.A.; Guenther, P.L.; Stedman, D.H.; Stephens, R.D.; Groblicki, P.J.; Johnson, B.J.; Huang, S.C. "On-road remote sensing of carbon monoxide and hydrocarbon emissions during several vehicle operating conditions", presented at Air & Waste Management Association International Specialty Conference on PM10 Standards and Non-traditional Source Controls, Phoenix, AZ 1992.
17. Stedman, D.H.; Bishop, G.A. "Remote sensing for mobile source CO emission reduction," U.S. Environmental Protection Agency, Las Vegas, 1991; EPA 600/4-90/032.
18. Lawson, D.R.; Groblicki, P.J.; Stedman, D.H.; Bishop, G.A.; Guenther, P.L. "Emissions from in-use motor vehicles in Los Angeles: A pilot study of remote sensing and the inspection and maintenance program," *J. Air &Waste Manage. Assoc.* **1990**, *40*, 1096.
19. Federal Register, 21(60), Part II, 1966.
20. Federal Register, 33(108), Part II, 1968.
21. Federal Register, 35(214), Part II, 1970.
22. Federal Register, 36(128), Part II, 1971.
23. Berg, D. "Survey of sources of test variability in the 1975 federal test procedure", U.S. Environmental Protection Agency, Motor Vehicle Emissions Laboratory, Ann Arbor, MI., 1978.
24. Knepper, J.C.; Koehl, W.J.; Benson, J.D.; Burns, V.R.; Gorse, Jr., R.A.; Hochhauser, A.M.; Leppard, W.R.; Rapp, L.A.; Reuter, R.M. "Fuel effects in auto/oil high emitting vehicles," Soc. Automotive Eng.; Paper No. 930137; 1993.
25. Clean Air Act as Amended. "Provisions for attainment and maintenance of national ambient air quality standards," U.S. Government Printing Office, Washington, D.C., November 15, 1990; Public Law 101-549.
26. California Air Resources Board. "Technologies to improve the detection of high- emitting vehicles in a vehicle inspection program," El Monte, CA, December 1992.
27. Harrington, W.; McConnell, V. "Modeling in-use vehicle emissions and the effects of inspection and maintenance programs," *J. Air & Waste Manage. Assoc.* **1994**, *44*, 791-199.
28. United States General Accounting Office. "Air pollution: Unresolved issues may hamper success of EPA's proposed emissions program", report to the Chairman, Subcommittee on Oversight and Investigations, Committee on Energy and Commerce, House of Representatives, Washington, D.C., 1992; Paper No. GAO/RCED-92-288.
29. Smith, L.R. "Variability of I/M test scores over time," U.S. Environmental Protection Agency, Ann Arbor, 1988; EPA 460/3-88-008.
30. Stedman, D.H.; Bishop, G.A. "An analysis of on-road remote sensing as a tool for automobile emissions control," Illinois Department of Energy and Natural Resources, Springfield, IL, 1990; ILENR/RE-AQ-90/05.
31. Stedman, D.H.; Bishop, G.A.; Peterson, J.E.; Guenther, P.L.; McVey, I.F.; Beaton, S.P. "On-road carbon monoxide and hydrocarbon remote sensing in the Chicago area," Illinois Department of Energy and Natural Resources, Springfield, 1991; ILENR/RE-AQ-91/14.
32. California I/M Review Committee. "Evaluation of the California Smog Check program and recommendations for program improvements," Fourth Report to the Legislature, Sacramento, CA, 1993.
33. Harrington, W.; McConnell, V.D. "Cost effectiveness of remote sensing of vehicle emissions," in *Cost Effective Control of Urban Smog* (R.F. Kosobud, W.A. Testa, D.A. Hanson, Eds.), Federal Reserve Bank of Chicago, Chicago, 1993; pp. 53-75.
34. Colorado Department of Health. "1988-89 oxygenated fuels program," Final report to the Colorado Air Quality Control Commission, Denver, CO, 1989.
35. Guenther, P.; Stedman, D.; Bishop, G.; Lesko, J. "University of Denver carbon monoxide emissions control program," Final report to the City and County of Denver, CO, July 1994.
36. Pitchford, M.; Johnson, B. "Empirical model of vehicle emissions," *Environ. Sci. & Tech.* **1993**, *27*, 741.
37. Lawson, D.R. "The cost of M in I/M - Reflections on inspection/maintenance programs," *J. Air &Waste Manage. Assoc.* **1995**, *45*, 465.

About the Authors

Gary A. Bishop, Ph.D. (corresponding author) is a Research Scientist and Donald H. Stedman, Ph.D. is the Brainerd Phillipson Professor of Chemistry at the University of Denver, Chemistry Department, Denver, CO 80208-0179. Lowell Ashbaugh is an Associate Research Ecologist (PhD) at the University of California at Davis, Davis, CA.