PART B OF THE SUPPORTING STATEMENT

Determine Percentage of High Evaporative Emissions
Vehicles in the On-Road Fleet

OMB Control Number 2060-NEW

EPA ICR Number 2292.01

1 SURVEY OBJECTIVES, KEY VARIABLES AND OTHER PRELIMINARIES

Evaporative emissions from gasoline vehicles have been evaluated and regulated since the early 1970s. Gasoline vehicles have evaporative emissions control systems that control excessive evaporative emissions, which are essentially gasoline vapors. When these systems or the gasoline delivery system of a vehicle malfunction, excessive evaporative emissions can be emitted. The mass of evaporative emissions from individual vehicles has been quantified in previous studies [1, 2, 3], but the frequency of vehicles in such a state in the general population has been estimated based on limited data [1, 3, 4].

1(a) Survey Objectives

This data collection effort is a survey designed to estimate the percentage of high evaporative emissions vehicles in the on-road fleet of gasoline-powered passenger cars and light-duty trucks. Specifically, the primary question of this study is:

What fraction of the fleet is made up of high evaporative emissions vehicles?

Because evaporative emissions are made up of various types (diurnal, hotsoak, running losses, resting losses, fugitives, and gross liquid leaks) and because there is not a clear, well-accepted definition of "high evaporative emitter," the question can be re-stated in terms of two other questions:

a) What field method can serve as a practical and substantially accurate method of identifying high-emitting evaporative emissions vehicles (High Evaps)?
b) What fraction of the fleet is made up of High Evaps as defined by the above field method?

Official measurements of the evaporative emissions of a vehicle are made in a laboratory Sealed Housing Evaporative Determination (SHED) enclosure. However, those measurements are expensive and require several days for each vehicle tested. This study requires a field method that is inexpensive and fast and whose results are correlated with official lab SHED results. Based on our cumulative experience, we believe that high-emitting evaporative emissions vehicles can be identified by the high levels of ambient hydrocarbon compounds (HC) that build up in a portable SHED after a vehicle is enclosed and its engine is turned off. We expect that this technique can be the field method to answer the question: "Is this vehicle a High Evap?" By using the portable SHED to measure the evaporative emissions of a stratified sample of vehicles in the in-use fleet, the fraction of High Evaps in the fleet can be estimated.

1(b) Key Variables

Variables to be surveyed or measured include:

- Vehicle identifiers: License plate and Vehicle Identification Number
- Vehicle description : Model year, make, model, and odometer reading
- Vehicle usage and maintenance history through a vehicle owner survey

- Time trace of HC concentration of the air inside the portable SHED after a vehicle's engine is turned off and the portable SHED doors are sealed.

- Measured values of high evaporative emissions vehicle screening methods:

  - Remote Sensing Device HC measurement
  - Modified California Method (Under-hood visual inspection and electronic HC vapor leak detector inspection)
  - Infrared video camera

1(c) Statistical Approach

We have selected a statistical approach for this effort for two reasons:

- While a census or partial census would be ideal, the effort and expense required is prohibitive.
- To meet the objectives for use of these data, it is necessary to draw valid and defensible inferences from sets of equipment surveyed or measured to equipment populations at wider scales, such as the county or state or nation. This requirement in itself rules out non-probabilistic approaches.

1(d) Feasibility

Obstacles to Participation. We do not anticipate substantial obstacles to participation. We have planned to conduct this survey in an area where passenger cars and trucks conduct normal everyday business like a shopping mall or a state vehicle inspection station. As vehicles come into such an area for conducting their business, vehicles will be first screened in a non-intrusive way to identify possible high evaporative emissions vehicles. A sample of these vehicles will be approached for some additional testing as they park their vehicles or wait in line. Vehicle drivers will be requested to participate in a questionnaire about their experience of using their vehicle, and their vehicle characteristics will be noted.

Availability of Funds. At present we expect to have adequate funds available to conduct the survey as designed. Funds will be contributed by two government partners and one industry partner. The first government partner is the Assessment & Standards Division within the EPA Office of Transportation and Air Quality (OTAQ). The industry partner is the Coordinating Research Council (CRC), a nonprofit research organization whose members include the American Petroleum Institute (API), the Society of Automotive Engineers (SAE), General Motors, Ford Motor, Chrysler, Volkswagen and Honda. However, if funding shortfalls occur, we can take measures to reduce sampling costs. One possibility would be to reduce the number of vehicles in the study.

2 SURVEY DESIGN

2(a) Target Population and Coverage

The target population includes all gasoline-powered passenger cars and light-duty trucks. Passenger cars are light-duty vehicles with gross vehicle weight ratings of less than 6000 lbs. Light-duty trucks are trucks with gross vehicle weight ratings of less than 8500 lbs. Passenger cars and light-duty trucks form the majority of the on-road motor vehicle fleet. Nationally, they account for 96.6% of the on-road vehicle fleet and 89.0% of the total on-road vehicle miles traveled. Heavy-duty vehicles account for the remainder of the on-road vehicle fleet and the on-road vehicle miles traveled.

At this point, work is anticipated in two geographical areas. The first area, which will be used for pre-testing and the pilot study, is Denver, Colorado. Denver was chosen for several reasons. The Colorado Department of Public Health and Environment (DPHE), which is located in Denver, operates a laboratory SHED, runs the state's inspection/maintenance (I/M) program, and runs the state's on-road RSD measurement program. Those DPHE resources will be used in the pilot study. An I/M inspection station will be used as a convenient source of private vehicles to solicit for the pilot study. Finally, the team that originally developed the RSD technique is located at the University of Denver and will participate in the pre-testing activities. The Denver survey should be completed in the summer of 2008.

The second area, which will be used for the main study, will be an area that does not have an I/M program. We expect that a non-I/M area fleet will have more evaporative emissions problems than an I/M area fleet. A non-I/M area should provide an upper bound on the fraction of High Evaps in a fleet and therefore should represent the worst case High Evap scenario. The area chosen will likely be in the southern United States where the climate is still warm later in the year to enhance evaporative emissions.

2(b) Sample Design

2(b)(i) Sample Frame

The sample frame will include all vehicles that have received screened RSD measurements. All of these vehicles have an opportunity to be selected for participation in the study. Registration data will be used to see if the vehicles with screened RSD measurements are representative of the vehicles registered in the study area fleet and in the national fleet. EPA will provide the model year distribution of the national fleet, and state registration data will reveal the model year distribution of the study area fleet. The model year distribution of vehicles screened by RSD – as determined by the license plates via connection to the state registration database – will be compared to the model year distribution of vehicles registered in the study area and in the national fleet.

2(b)(ii) Stratification Variables

Two stratification variables will be used in this study:

1) RSD HC, and

2) Vehicle Age.

RSD HC. Since most vehicles are not excessive emitters of evaporative emissions, stratification based on RSD HC will be used to enrich the sample with vehicles that are potentially high evaporative emitting. In the survey we will use remote sensing to screen vehicles into nine RSD HC categories. As described below, we have circumstantial evidence that the HC channel of certain types of RSD instruments may be sensitive to vehicles with high evaporative emissions levels. By preferentially sampling more vehicles from the higher RSD HC bins, we believe that the effort can capture a larger fraction of High Evaps in the 1000-vehicle survey than could be captured by completely randomized sampling from the fleet as a whole.

Recent data collected in the California RSD Pilot project [5] suggest that the tailpipe HC channel of the RSD instrument used in that study, the ESP Accuscan 4000, may be influenced by a vehicle's evaporative emissions, which are HCs. Here is the evidence.

The RSD instrument uses a light beam shining across the roadway to measure pollutants in a vehicle's tailpipe plume. The instrument has HC, CO, and NO channels. In the California study, a few days to several months after vehicles were measured by the on-road RSD instrument, a subset of the vehicles naturally received their regular state inspection program tailpipe emissions test, which is known as the Acceleration Simulation Mode (ASM) test. Analysis of bins of the 76,982 paired RSD and ASM results showed a quite linear relationship for CO and NO when the logit of the ASM failure rate was plotted against the natural log of the RSD concentration. Figures 1 and 2 show the relationships for CO and NO. Straight line fits of the trends and 95% confidence limits on the individual points are included. The upward trend in both plots shows that, on the average, vehicles that have higher measured RSD tailpipe concentrations were more likely to fail their state tailpipe emissions inspection for the same pollutant.

Figure 1. ASM CO vs. RSD CO

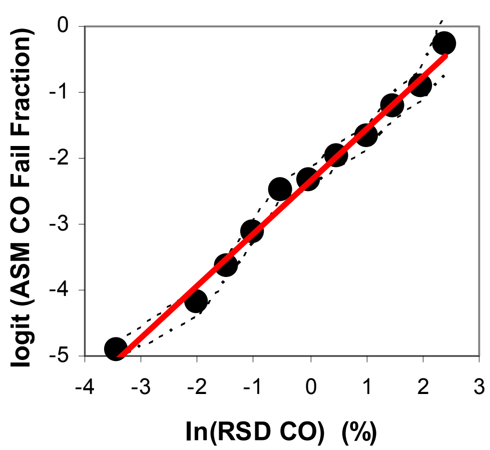


Figure 2. ASM NO vs. RSD NO

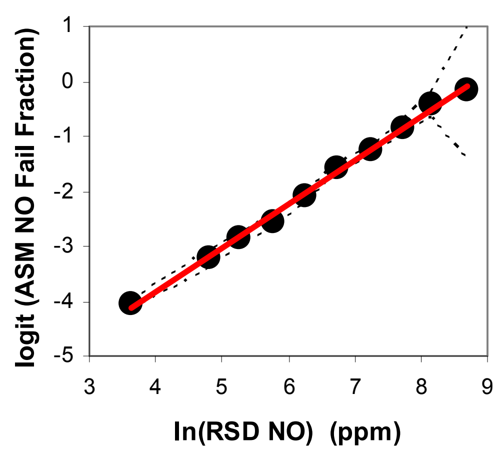
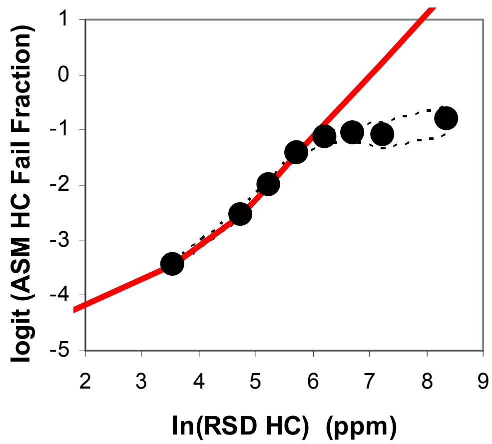


Figure 3. ASM HC vs. RSD HC



However, when the same type of plot is made for the analysis of bins of the paired data for HC, two different regions of behavior are observed. See Figure 3. In the low RSD HC region (ln RSD HC < 6), the ASM HC fail rate follows the trends seen for CO and NO. That is, vehicles with higher and higher on-road RSD HC concentrations are more and more likely to fail the inspection station ASM HC tailpipe test. The data fall on a relatively linear trend as approximated by the solid red line. However, on the right side of the plot (ln RSD HC > 6), a second, different trend is observed. Here, the data points reach a plateau (logit ASM HC fail fraction = -1), which means that about 27% of the vehicles fail the tailpipe ASM HC tailpipe test even though their on-road RSD HC concentrations range from 400 ppm up to 4000 ppm. That is, in the high RSD HC region, ever increasing RSD HC concentrations do not translate into an ever increasing probability of failing the inspection station ASM HC tailpipe test.

One explanation, but perhaps not the only explanation, for the observed HC behavior is the presence of High Evaps in the fleet sample. High Evaps could pass the inspection station ASM HC test because the ASM test is a tailpipe test, and therefore it is not influenced by evaporative emissions, which are emitted only from the fuel handling systems of vehicles – not from their tailpipes. However, when a vehicle drives on the road, evaporative emissions can become mixed with tailpipe emissions in the plume behind the vehicle. Depending on how the RSD instrument processes the data obtained from its light beam, evaporative emissions could increase the reported RSD HC readings over what one would expect on the basis of tailpipe emissions alone. If the evaporative emissions are very high, the increase could be large enough to cause points on the linear trend in Figure 3 to be moved to the right of the expected trend depicted by the red line.

Since evaporative emissions testing of vehicles was not performed in the California RSD Pilot study, this explanation of the trends in Figure 3 is unconfirmed. Nevertheless, the explanation makes sense. In addition, Don Stedman, the developer of the RSD technique, is familiar with the data processing algorithm of the Accuscan-4000 instrument and believes that its high RSD HC readings may be influenced by High Evaps. Since algorithms of other RSD instruments may be less sensitive to evaporative emissions, this finding, if confirmed, could lead to the development of new RSD processing algorithms that could specifically target the on-road measurement of evaporative emissions.

Vehicle Age. Based on the estimates by Landman [4], we expect that High Evaps are more prevalent in older vehicles. We will allocate the sample equally among the age strata because the age trend is too uncertain to set up an optimized allocation. We will use three strata within each RSD HC stratum to ensure that equal numbers of vehicles within each RSD HC stratum are obtained for three age strata: the oldest one-third, the middle one-third, and the newest one-third of vehicles in each RSD HC stratum. This sampling method for vehicle age will guarantee that the vehicles sampled within an RSD HC bin will have a range of ages and be representative of the ages in the fleet. The two age cutpoints for each RSD HC stratum will be determined by an analysis of existing RSD HC data from California or Virginia.

2(b)(iii) Sample Sizes

The goal is to select about 1000 vehicles for the determination of their evaporative emissions with the portable SHED and for the evaluation of the three high-evaporative emissions screening methods. Because, as discussed above, we have evidence that RSD HC may be influenced by evaporative emissions, we plan to select the 1000 survey vehicles from the fleet using RSD as a screening tool. We know from past experience that remote sensing can screen many more vehicles than the required 1000 vehicles in a relatively short time. Therefore, using remote sensing to select survey vehicles is feasible.

Designing a stratified sampling strategy that can achieve the desired precision (see Section 2(c)(i)) requires an estimate of the abundance of High Evaps in the fleet as a function of the stratifying variable. Unfortunately, no dataset exists that can clearly define High Evap levels. However, if we assume that the trends in Figure 3 are caused by High Evaps, then the data collected in the California study can be used to develop a stratified sampling design to meet the precision target for this study.

Table 1 summarizes the data from that study that will be used to develop a stratified sampling plan for this study. Columns B and C give the bin definitions in terms of the RSD HC as measured by the Accuscan-4000 instrument. Columns D and E give the distribution of vehicles that had an inspection station ASM test that followed the on-road RSD measurement. These columns indicate that the distribution of RSD HC emissions in the fleet is highly skewed with about 83% of the vehicles having RSD HC emissions below 148 ppm. Column F shows the number of High Evaps that we estimated from Figure 3. Specifically, we assumed that the gap between the data points and the red line in Figure 3 is caused by the presence of High Evaps in the fleet. This is estimated to be 0.75% of the vehicles. Column G gives the estimated probability (=F/D) that vehicles in each bin are High Evaps.

Within each RSD HC bin, vehicles will be selected with three sub-quotas for model years: older vehicles, average-age vehicles, and newer vehicles. Each sub-stratum will have a quota of 33% of the required count for the RSD HC bin. The sub-strata definitions will be determined by an analysis of California or Virginia RSD data.

From the information in Columns A through G, we used the technique for optimum stratified sampling, which is described in Appendix A, to develop a plan for achieving the precision target. This optimized stratified design minimizes the uncertainty in the estimated fraction of High Evaps in the fleet. We found that a stratified random sample of 858 vehicles following the quotas in Column H would produce a half width of the 95% confidence interval of 25% expressed as a percent of the fraction of High Evaps in the fleet. That is, the 95% confidence interval would be 0.75% ± 0.19%. Because the quotas in Column H call for "partial" vehicles, the real sampling plan, which of course must sample whole vehicles, will be a close approximation of the optimal design. Consequently, the resulting mean fraction of High Evaps and its uncertainty will be slightly different from the values calculated by the optimum design in Table 1.

Table 1. Optimum Stratified Random Design for High Evap Fraction

A	B	C	D	E	F	G	H
Bin Definitions: RSD HC (ppm)			Vehicle Population		High Evap Vehicles (estimated)		Optimum Stratified Random Design
Bin	Greater Than	Less Than or Equal To	Number of Vehicles in Bin	Population Fraction in Bin (W_h)	Number of High Evaps in Bin	Fraction of Vehicles in Bin that are High Evaps (p_h)	Number of Optimum Design Vehicles in Bin (N_h)
1	-Inf	0	24101	0.3131	0	0.000002	13.98
2	0	90	32959	0.4281	1	0.00005	85.66
3	90	148	7094	0.0922	3	0.0005	58.20
4	148	245	5476	0.0711	15	0.0027	110.25
5	245	403	3644	0.0473	49	0.013	161.54
6	403	665	2138	0.0278	107	0.050	179.48
7	665	1097	917	0.0119	143	0.156	128.36
8	1097	1808	365	0.0047	112	0.308	65.00
9	1808	Inf	288	0.0037	148	0.513	55.53
All			76982	1.0000	578	0.75%	858.00

There are at least two reasons that the rate of occurrence of High Evaps in each of the nine bins may not turn out to be the expected values shown in Column G of Table 1. First, our method to determine the rates is based on our interpretation of the trends seen in Figure 3 in terms of evaporative emissions. Second, while RSD can estimate the average emissions of a fleet of vehicles, RSD's ability to properly classify individual vehicles in emissions bins is subject to considerable uncertainty caused by the variability of vehicle emissions and by variability in the RSD technique itself.

While the stratified plan in Table 1 is optimum for the information given in Columns A through G, if the observed fraction of High Evaps in each bin is not as expected – for whatever reason – the precision of the calculated fraction of High Evaps will suffer. An analysis of the design shows precision is especially sensitive to the number of High Evaps in Bin 1, which contains 31% of the vehicles in the fleet but from which only 14 vehicles would be sampled. For example, if instead of getting the expected zero High Evaps in Bin 1, the study finds one High Evap there, the 95% confidence interval would be 2.96% ± 4.23%, which is a half width of the 95% confidence interval of 143% expressed as a percent of the fraction of High Evaps in the fleet. The 143% precision is far above the precision target of 25%.

To reduce the influence of a few "surprise" results in the lower RSD HC bins on the precision, it is prudent to supplement the optimum stratified design with additional vehicles to be sampled in the lower bins. As shown in Table 2, we have chosen to add 142 vehicles to the optimum design's lowest three bins, so that each of those bins will have 100 vehicles sampled. Now, if Bin 1 has 7.14 High Evaps of its 100 vehicles (this is equivalent to 1 in 14), the 95% confidence interval would be 2.96% ± 1.59%, which is a half width of the 95% confidence interval of 54% expressed as a percent of the fraction of High Evaps in the fleet. The 54% is still elevated above the precision target of 22% but not as severely elevated as when only 14 vehicles were sampled in Bin 1. On the other hand, if Bin 1 turns out to have zero High Evaps as expected and as shown in Column G, the precision target of 25% will still be met.

Increasing the counts allocated to the three lowest bins takes care of the greatest vulnerability regarding results that differ significantly from what is expected. Larger-than-expected observed counts could occur in those bins by random chance, even if the information on which the optimal stratified plan were correct.

Above, we mention that the stratified sampling plan presented in Table 1 is optimal for the information presented in Columns A through G. If the information is other than that shown in the table, the experimental design may not be optimal. However, oversampling the high emitters relative to their representation in the fleet moves things in the right direction, even if the design is not exactly optimal. Again, we have addressed the greatest vulnerability to suboptimality by increasing the counts in the lowest bins.

Further, assuming the historical information about the fraction of the fleet that falls in each bin (Column E in Table 1) is correct, the estimate of the percent of high emitters for the fleet will be unbiased in any case. This follows from the way the experimental results for individual strata are combined to produce the fleet estimate; the details are given in Appendix A. This unbiased property does not depend on the initial expectation regarding the fraction of high evaporative emitters in the individual bins based on the discussion accompanying Figure 3.

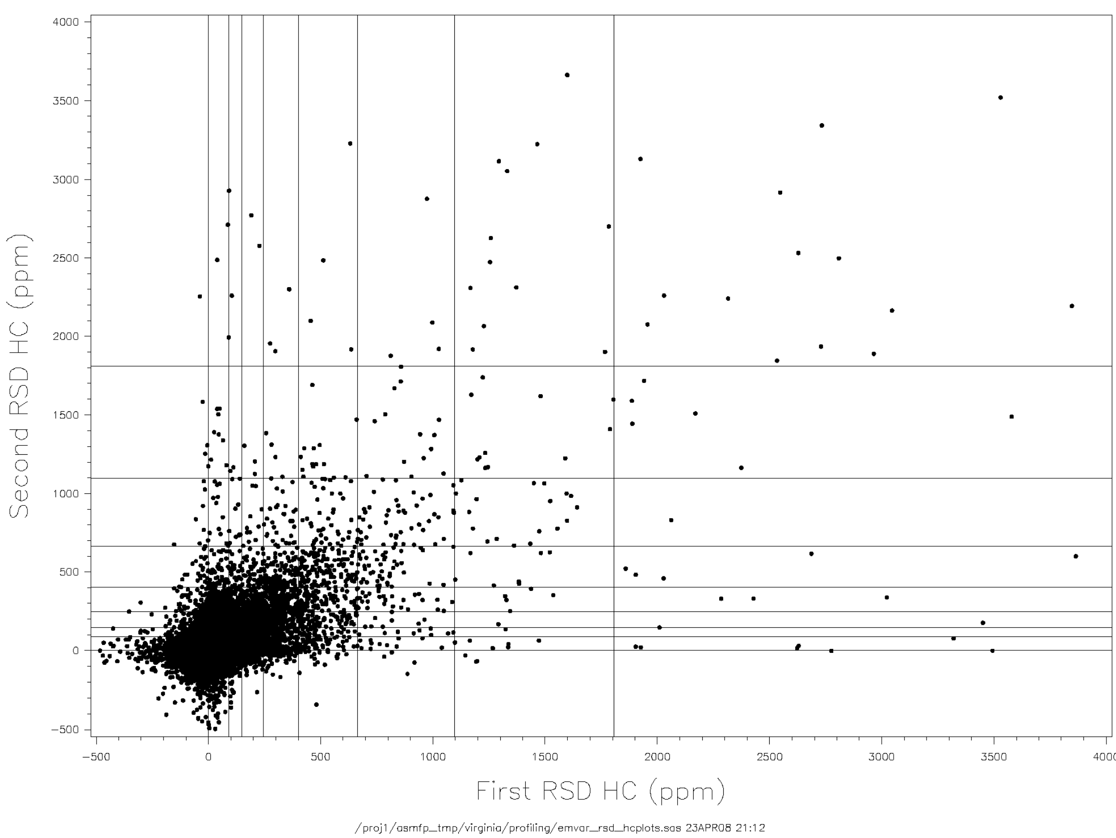
Table 2. Modified Stratified Random Design for High Evap Fraction

A	B	C	D	E	F	G
Bin Definitions: RSD HC (ppm)			**Optimum Stratified Random Design**	**Modified Optimum Stratified Random Design**		
Bin	**Greater Than**	**Less Than or Equal To**	**Number of Optimum Design Vehicles in Bin (N_h)**	**Number of Additional Vehicles in Bin**	**Total Number of Vehicles to be Sampled in Bin**	**Expected Number of High Evaps in Bin**
1	-Inf	0	14	86	100	0
2	0	90	86	14	100	0
3	90	148	58	42	100	0
4	148	245	110	0	110	0
5	245	403	162	0	162	2
6	403	665	179	0	179	9
7	665	1097	128	0	128	20
8	1097	1808	65	0	65	20
9	1808	Inf	56	0	56	29
All			858	142	1000	80

Vehicles will be screened for possible participation in the study based on their first RSD HC emissions measurement. However, RSD HC measurements are subject to large variability. Therefore, because of large expected regression-toward-the-mean effects, it is important to recognize that the RSD HC bin assignments and all data analyses must be based on a second RSD reading. Because the set of vehicles will be selected by RSD HC, a measurement subject to error, the second RSD HC measurements, on which the results of the study will based, will, in many cases, be different from the first RSD HC measurement. These differences will affect the number of vehicles that will need to be scanned by the first RSD measurement in order to fill the RSD HC bins according to the second RSD measurement.

In 2002 Virginia conducted a large RSD pilot study using Accuscan-4000 instruments [6]. Data from that study demonstrates how the degree of variability in RSD HC measurements affects the number of vehicles assigned to RSD HC bins. Figure 4 compares duplicate RSD HC measurements in the Virginia study. For this figure, duplicate RSDs are defined as two RSD measurements taken by the same RSD instrument at the same roadway location within two days of each other. This data can be used to simulate the distribution of second RSD HC measurements that would be obtained when vehicles are selected according to their first RSD HC measurement.

Figure 4. Duplicate RSD HC Measurements from Virginia



The figure shows that only a portion (45%) of vehicles that fall into Bin 9 (RSD HC > 1808 ppm) by the first RSD measurement would remain in Bin 9 for the second RSD measurement. Since the second RSD measurements are the values that will be used to analyze the study's data, the number of vehicles assigned to Bin 9 would be low. The same effect is seen to be true for Bin 8; only a portion (14%) of the vehicles screened as Bin 8 vehicles would remain in Bin 8 for the second RSD measurement. However, several of the vehicles screened as Bin 8 vehicles would actually be assigned to Bin 9 by the second RSD measurement, and this helps bring up the number of vehicles assigned to Bin 9 by the second RSD measurement. This trend continues for all of the bins.

Table 3 shows a simulated result (the average of 3000 realizations) using the Virginia dataset of RSD duplicate measurements if vehicles are selected strictly by the first RSD HCs (Column D) during screening. The resulting average bin quotas for the stratified sample as assigned by the second RSD HC measurement would not be met as seen in Column E. Specifically, the lower RSD HC bins would be over-populated and the upper RSD HC bins would be under-populated relative to the desired quotas. Therefore, we need to have a strategy to compensate for this regression-toward-the-mean effect. The strategy is to use screening RSD HC bin quotas that are lower than needed for Bins 1 and 2 and higher than needed for Bins 4 through 9. Additionally, the first RSD HC quotas need to be selected so that <u>for a single realization</u> the probability is high that the number of vehicles in each of the second RSD HC bins is greater than or equal to the desired quotas. Otherwise, the precision target will probably not be achieved in the field. To take these effects into account, the screening quotas for the first RSD HC bins will be calculated based on the Virginia duplicate RSD data.

Table 3. Simulated Stratified Screening RSDs and Measurement RSDs

A	B	C	D	E
Bin Definitions: RSD HC (ppm)			**Select RSD#1 such that RSD#1 has Desired Quotas**	
Bin	**Greater Than**	**Less Than or Equal To**	**Vehicles Selected Using *Screening* RSD Bin (RSD #1)**	**Vehicles Observed in *Measurement* RSD Bin (RSD #2)**
1	-Inf	0	100	163
2	0	90	100	236
3	90	148	100	92
4	148	245	110	111
5	245	403	162	126
6	403	665	179	122
7	665	1097	128	74
8	1097	1808	65	33
9	1808	Inf	56	44

2(b)(iv) Sampling Methods

Vehicles will be sampled using a stratified random design. The sample will be stratified with optimal allocation among RSD HC emissions bins and, within each of those bins, equally allocated among vehicle age bins. Simple random sampling will be used for the RSD-by-Age strata.

2(c) Precision Requirements

2(c)(i) Precision Targets

The main quantity to be determined in this study is the fraction of the fleet vehicles that are high evaporative emitters (High Evaps). We would like to know with high confidence that this quantity has an uncertainty of no more than ±25% of the value. Specifically, the precision target is that the half width of the 95% confidence interval of the fraction is no more than 25% of the fraction.

2(c)(ii) Non-Sampling Error

2(c)(ii)(1) Frame-coverage error

This error is defined as potential bias in key variables resulting from imperfections in the sample frame. The central issue is incomplete coverage, in which members of the target population are simply absent from the frame. The bias that may result from incomplete coverage may reduce the representativeness of the sample in a way analogous to that from whole-survey non-response. We have incorporated measures in the survey plan to detect and reduce the effects of these errors on the survey results.

2(c)(ii)(2) Non-response error

As in any survey, non-response is one of the most important potential sources of error in final results. Survey non-response occurs when no response at all is obtained from a potential participant in the study, whereas item-nonresponse occurs when a respondent provides responses to some but not all items. Survey non-response occurs if a respondent refuses to participate. Item-nonresponse may occur in a number of ways. A respondent may answer some items but refuse others, or may break off an interview for unrelated reasons. A form of item-nonresponse detrimental to emissions measurement but unrelated to the respondent could occur in cases where equipment malfunction or measurement errors make emissions datasets unsuitable for subsequent analysis.

2(c)(ii)(3) Measurement error

The measurement of emissions during normal equipment operation requires the use of complex instrumentation in a harsh environment. The emissions measurement instrument is specifically designed to collect data during normal operation. Additional steps will be take prior to and following data collection to detect measurement errors in resulting data.

Calibration. All instruments including the remote sensor, handheld HC detection device, HC instrument in the portable SHED will be calibrated regularly.

2(c)(ii)(4) Equipment malfunction

Following the measurements based on the various instruments, quality assurance measures will be undertaken to verify that the instruments operated correctly and that the results are reliable for further analysis. The QA process will involve the use of computer programs that automatically scan the time-series for patterns that may suggest instrument error, combined with graphic presentation of the data to allow case-by-case visual inspection.

2(c)(ii)(5) Respondent error

The emphasis on collection of key information for the survey through direct inspection and instrumentation involves a conscious decision to reduce reliance on human memory to the maximum extent possible. A primary example is the use of electronic dataloggers to measure vehicle emissions. As much as possible, we have restricted interview items to general questions that can be easily answered without involved or detailed estimation and without heavy reliance on human memory.

2(c)(ii)(6) Data entry error

Emissions results and other data collected electronically will not be input manually. Data files will be downloaded directly from the measurement instrument and transferred to the database, following quality-assurance procedures.

2(d) Measurement Design

2(d)(i) Screening Using Remote Sensing Device

All vehicles will have the emissions plume scanned by an RSD instrument to measure emissions concentrations. RSD instruments perform these measurements by shining a light beam across the roadway. Associated equipment will also simultaneously determine other quantities. All of these measurements will be performed without notifying the vehicle driver that they are being taken. For each vehicle the following quantities will automatically be taken as the vehicle passes by the RSD instrument:

- Item 1: DateTime: The date, hour, minute, and second of the RSD measurement.
- Item 2: Speed and Acceleration: The speed and acceleration of the vehicle.
- Item 3: RSD Emissions: The concentrations of HC, CO, and NO in the vehicle's plume.
- Item 4: License Plate: A digital image of the rear of the vehicle so that the license plate may be determined.

2(d)(ii) Solicitation

Based on the screening RSD HC values and the number of vehicles desired for each RSD/age bin in the stratified random plan, a sample of passenger cars and light-duty trucks will be approached by the solicitor. The solicitor will:

- Item 1: Offer incentive: The solicitor will offer a $20 incentive whether or not the vehicle owner participates.

- Item 2: Introduction: Explain that an emissions study is being conducted. Explain that a mechanic would like to perform measurements that would take about one hour. The solicitor will explain that the measurements would involve driving the car past the RSD unit, performing under-hood inspections, and testing the air in the portable SHED after the vehicle has sat in it for a few minutes.

- Item 3: Ownership: Ask if the driver owns the vehicle. Only vehicles with their owners driving will be eligible for participation.

- Item 4: Model Year: Verify the model year of the vehicle.

2(d)(iii) Evaporative Emissions Testing

The vehicles whose owners agree to participate would then undergo the following tests:

- Item 1: RSD Emissions: A technician will drive the vehicle past the RSD unit. The same type of data will be recorded as for the screen drive as described in 2(d)(i).

- Item 2: Portable SHED Emissions: The vehicle will be driven into a small tent with non-permeable walls. After closing the tent, turning off the vehicle's engine, and waiting a short time, the HC emission concentration of the air inside the tent will be measured. This will serve as an estimate of the true evaporative emissions of the vehicle. Following the measurement, the vehicle will be driven out of the tent, and the air in the tent will be vented.

- Item 3: Look and smell inspection: A visual and olfactory inspection of evaporative emission control systems to look for missing, malfunctioning, damaged, or disconnected components.

- Item 4: Handheld electronic HC sniffer inspection: Detection of high evaporative emissions using a handheld HC vapor detector. The small probe of the detector will be moved around components, fittings, and hoses to try to find escaping HC vapors.

- Item 5: Infrared Video Camera: The infrared video camera will record video of those areas around the vehicle where evaporative emissions might be present. This will include over the engine compartment with the hood opened and closed, and around the gasoline fill pipe.

3 PRETESTS AND PILOT TESTS

3(a) Pretests

The RSD instruments, portable SHED, handheld HC sniffer, and infrared video camera will be pre-tested by making measurements on surrogate evaporative emissions. The surrogate emissions will be artificially produced by various means including leaving off the gas cap of the vehicle and placing a gasoline-soaked rag under the hood of a test vehicle.

We will perform a cognitive test of the vehicle owner questionnaire. We will administer the questionnaire to a small sample of actual vehicle owners. Then, we will ask the owners questions to determine if the owners understood and how they received the questions in the questionnaire. Any misunderstandings that are revealed will be addressed by re-wording the questionnaire to improve it.

3(b) Pilot Tests

3(b)(i) Pilot Test Objectives

For the purposes of refining procedures for the main study, we will conduct a pilot study. The following questions will be answered by the pilot study:

a) How well do the results of the portable SHED method agree with the results of the standard laboratory SHED method? What are the characteristics of portable SHED testing for measuring the evaporative emissions of vehicles in the field? This includes issues such as ease of implementation, cost, number of vehicles tested in the portable SHED per day, level of personnel necessary, and measurement precision and accuracy.

b) What is the approximate fraction of the fleet that is made up of high evaporative emissions vehicles?

c) Can portable SHED measurements be performed in a way such that the results can be used to estimate the distribution of evaporative emissions of the fleet?

d) What are the characteristics of the three methods (RSD, modified California method, infrared video camera) for screening vehicles as High Evaps?

- What are the accuracy characteristics (four-quadrant, true-positive, and false-positive) of the screening methods for identifying high evaporative emitters?
- What are the practical characteristics of the screening methods? This includes issues such as ease of implementation, cost, time to complete one test, level of personnel necessary.

e) What refinements does the pilot suggest for the design of the larger study?

f) Based on the experience of the pilot study, what refinements or modifications would be considered for the design of a larger study?

3(b)(ii) Pilot Test Design

A pilot test will be run in the Denver area to test vehicle screening and solicitation procedures and field evaporative emissions measurement techniques and to obtain initial measurements of the occurrence of High Evaps in the nine RSD HC bins. The sampling design for the pilot study is 10% of the Table 2 design.

The pilot design is shown in Table 4. Of the 101 vehicles to be sampled in the pilot, 8 vehicles are expected to be High Evaps as shown at the bottom of Column G. The primary High-Evap-occurrence results of the pilot study are the nine High Evap fractions shown in Column H. With these nine values and population fractions for any application population (such as the nationwide fleet), the overall fraction of High Evaps in the application population can be calculated as described in Appendix A and as demonstrated by Table 4. For demonstration purposes this sample calculation uses the California RSD Pilot study population fractions shown in italics in Column D. To estimate the High Evap fraction in an application population, the Column D values must be changed to the actual population fractions for the application fleet in question.

Table 4. De-Stratification for the Pilot Study Results

A	B	C	D	E	F	G	H	I
Bin Definitions: RSD HC (ppm)			***Application* Vehicle Population**	**Modified Optimum Stratified Random Design Results**			**De-Stratifying Calculations**	
Bin	**Greater Than**	**Less Than or Equal To**	**Population Fraction in Bin (W_h)**	**Size of Screening Sample Needed to Fill Bin**	**Total Number of Vehicles Sampled in Bin (N_h)**	**Number of High Evaps Found in Bin**	**Fraction of Vehicles in Bin that are High Evaps (p_h)**	**Std Dev of the High Evap Probability for this Bin (σ_h)**
1	-Inf	0	*0.3131*	32	10	0	0.000	0.0000
2	0	90	*0.4281*	23	10	0	0.000	0.0000
3	90	148	*0.0922*	109	10	0	0.000	0.0000
4	148	245	*0.0711*	155	11	0	0.000	0.0000
5	245	403	*0.0473*	338	16	0	0.000	0.0000
6	403	665	*0.0278*	648	18	1	0.056	0.2291
7	665	1097	*0.0119*	1091	13	2	0.154	0.3608
8	1097	1808	*0.0047*	1476	7	2	0.286	0.4518
9	1808	Inf	*0.0037*	1604	6	3	0.500	0.5000
All			1.0000		101	8		
							Estimated % High Evaps in Fleet:	**Half-Width of the 95% Confidence Interval:**
							0.66%	0.43%

Relative Uncertainty:	66%

The de-stratification results in Table 4 show that even with 101 vehicles sampled and 8 High Evaps detected, the design can determine the High Evap fraction of an application fleet that is near the actual fleet High Evap fraction. However, the table shows that the half width of the 95% confidence interval is expected to be 66% of the estimated fraction of High Evaps in the fleet. This precision value of 66% is considerably larger than the precision target of 25% that the 1000-vehicle design can achieve. Thus, as expected the lower number of vehicles sampled in the pilot results in poorer precision. The pilot gives the study a chance to actually measure the rates of High Evaps found in the nine RSD HC bins while expending only a portion of the study's resources. If the pilot results are substantially different from the results expected from the analysis of the California RSD Pilot data, then the stratified design for the 1000-vehicle test may need to be re-optimized.

3(b)(iii) Sampling for Measuring Portable SHED Precision and Accuracy

Because the main study will use the portable SHED to determine the "true" High Evap status of the 1000 vehicles selected for testing, we must determine its precision and accuracy in the pilot study so that we can be confident that the portable SHED can be relied upon for the main study.

Precision. A subset of the vehicles participating in the portable SHED measurements will receive duplicate portable SHED measurements. The vehicles receiving duplicate testing will be selected so that they span a range of portable SHED emissions results. This is necessary so that the precision can be estimated across the range of portable SHED emissions that are produced.

Accuracy. To estimate the accuracy of the portable SHED, a fraction of vehicles during the pilot will be requested to participate in laboratory SHED testing at a certified lab near the recruitment site. These data will be used to compare the laboratory result with the portable SHED result. Vehicles for the laboratory SHED testing will be selected so that there are High Evap vehicles and Low Evap vehicles as determined by the portable SHED. We would like to be at least 90% confident that a positive association exists between the portable SHED and the lab SHED. The following discusses how the results can be analyzed to determine the level of confidence.

A test can be performed to determine whether there is a statistically significant association between the high-emitter determinations by the portable SHED and by the lab SHED. The test is based on a 2-by-2 contingency table such as shown in Table 5. The test statistic is a chi-square statistic with one degree of freedom.

The cells along the main diagonal represent agreement (both methods indicate a high emitter, or both methods indicate a non-high emitter). The cells off the main diagonal indicate disagreement between the two methods.

Under the null hypothesis, there is no true association between the determinations by the two methods. If the difference between the two methods is zero within random variability, the chi-square value will equal its expected value (the number of degrees of freedom) within random variability. If the chi-square value is too large to be explained by random variability, we reject the null hypothesis and conclude that there is an association between the two types of measurements. Note that if any cell's expected value is less than or equal to 5, then Fisher's Exact Test should be used instead of the chi-square test.

Table 5. Simulated Results for Portable SHED Accuracy

Portable SHED	Laboratory SHED: Low Evap	Laboratory SHED: High Evap
Called High Evap	5	10
Called Low Evap	10	5

Following the example of Snedecor and Cochran [7], if the counts along the main diagonal are both 10 and the counts in both cells off the main diagonal are 5, the chi-square value is 3.33. The critical value for 90% confidence is 2.71. Thus, if we have 30 vehicles, equally split between high emitters and non-high emitters, and if the counts indicating agreement are twice the counts indicating disagreement, we have a statistically significant result with 90% confidence, with a little margin.

The results of the comparison between the portable SHED and the lab SHED depend not only on the total number of vehicles tested in both, but also on how the vehicles are distributed among the four quadrants in Table 5. In particular, it is important to select some vehicles that are Called High Evaps and some vehicles that are Called Low Evaps by the portable SHED.

In this study, the portable SHED will likely be used primarily as an index of the High Evap status of a vehicle (that is, each vehicle will be categorized as High Evap or Low Evap) and not necessarily as a measure of the *magnitude* of the mass of evaporative emissions of a vehicle. Nevertheless, the portable SHED can also be evaluated by comparing the mass evaporative emissions results obtained in the portable SHED and lab SHED. Those results can be compared using a scatter plot and correlation coefficient.

4 COLLECTION METHODS AND FOLLOW-UP

4(a) Collection Methods

Vehicle owner survey – A solicitor will collect recent vehicle usage and maintenance history information at the testing site by conducting a personal interview using questions 1 through 10 of the questionnaire in Appendix B.

Vehicle information – A technician will visually examine the vehicle to collect the vehicle identity information listed in Appendix C. Some items will be documented with a camera to reduce the chance of transcription errors.

Remote Sensing Device HC measurement – The concentration of HC in the plume behind a vehicle will be measured remotely using RSD instruments. The RSD instrument sends a light beam across the roadway. Immediately after a vehicle passes through the light beam, the RSD instrument reports the concentration of HC, CO, and NO in the vehicle's plume. Other associated equipment will measure the vehicle's speed and acceleration. A video camera will take an image of the rear of the vehicle so that the license plate can be used to determine the identity of the vehicle.

Modified California Method – A technician will make an inspection of evaporative emissions control system components and fuel system components on the vehicle. The vehicle will be inspected under the hood, around the fuel tank, and around fuel lines. The technician will perform a visual inspection, an olfactory inspection, and a gasoline liquid and vapor leak inspection using an electronic HC vapor detector.

Infrared video camera – A technician will use an infrared video camera to film areas under the hood, around the fuel tank, and around fuel lines.

Portable SHED – The time trace of HC concentration of the air inside the portable SHED will be obtained during a period (perhaps 15 minutes) after a vehicle's engine is turned off and the portable SHED doors are sealed. The actual procedure for measuring evaporative emissions in the portable SHED test will be determined during the development of the testing methods.

5 ANALYZING AND REPORTING SURVEY RESULTS

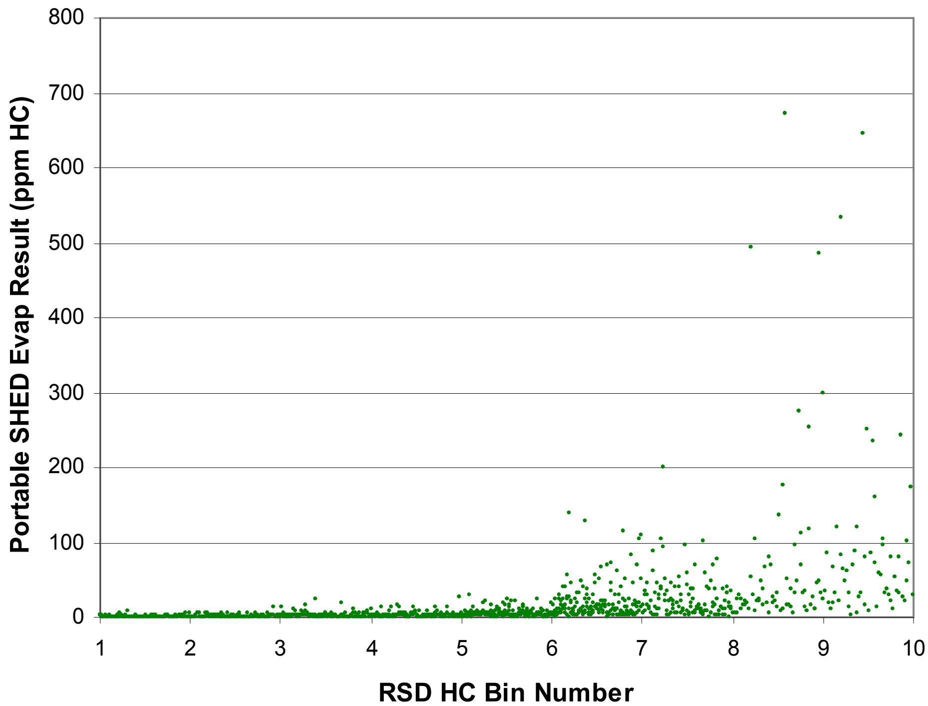
5(a) Data Preparation

5(a)(i) De-Stratifying the Results

Because the field tests will be performed following a stratified random design, the results need to be de-stratified to determine the measured characteristics on the basis of an application fleet. The application fleet is the fleet of interest for any particular calculation. For example, the application fleet may be the national fleet or the fleet of a particular state. De-stratification needs to be applied to the portable SHED evaporative emissions results to determine the High Evap fraction of the fleet and the distribution of evaporative emissions in the fleet, operational characteristics of the RSD High Evap vehicle selection method, and operational characteristics of the modified California High Evap vehicle selection method.

Figure 5 shows simulated results for the 1000-vehicle main study. The horizontal axis represents the RSD HC results in the nine bins. For example, the results of the 56 vehicles in Bin 9 have x-values between 9 and 10. Test results for vehicles with low RSD HC are on the left, and test results for vehicles with high RSD HC are on the right. Consequently, the figure has 1000 data points. The vertical axis represents the portable SHED evaporative emissions results. For the purposes of the simulation, we have manufactured specific portable SHED HC concentrations that are randomly selected from a separate log-normal distribution for each RSD HC bin. The log-normal emissions distributions were created so that the number of observations with concentrations above 50 ppm HC in each bin approximates the number of High Evap observations expected by Column G in Table 2.

Figure 5. Simulated Main Study Results



The results as plotted in Figure 5 can be used to determine calculated quantities for the analysis of an application fleet; however, they need to be de-stratified according the characteristics of the application fleet. To do this, the counts in each bin in Figure 5 need to be multiplied by a de-stratification factor so that the de-stratified counts in each bin are proportional to the number of vehicles that would be seen by RSD in each bin in the application fleet. The calculation of the de-stratification factors is shown in Table 6. We have assumed that the population fraction in each bin (Column D) is the same as the California RSD distribution. The sampling weights in Column F are calculated by dividing each bin's population fraction (Column D) by the number of vehicles sampled in each bin (Column E). Then, the de-stratifying factors (Column G) are calculated by dividing all sampling weights by the smallest sampling weight for all of the bins, which is 5.14 for Bin 9.

Table 6. Example De-Stratification Factors for the Portable SHED Results

A	B	C	D	E	F	G
Bin Definitions: RSD HC (ppm)			***Application* Population**			
Bin	**Greater Than**	**Less Than or Equal To**	**Number of Vehicles in Bin**	**Total Number of Vehicles Sampled in Bin (N_h)**	**Sampling Weights**	**De-Stratifying Factor**
1	-Inf	0	*24101*	100	241.01	46.86
2	0	90	*32959*	100	329.59	64.09
3	90	148	*7094*	100	70.94	13.79
4	148	245	*5476*	110	49.78	9.68
5	245	403	*3644*	162	22.49	4.37
6	403	665	*2138*	179	11.94	2.32
7	665	1097	*917*	128	7.16	1.39
8	1097	1808	*365*	65	5.62	1.09
9	1808	Inf	*288*	56	5.14	1.00

For the purposes of analysis of the portable SHED data, counts in any given RSD HC bin should be multiplied by the corresponding de-stratifying factor so that the results will approximate the results that would be obtained from a random sample of the application fleet.

5(b) Data Analysis

5(b)(i) Fraction of High Evaporative Emissions Vehicles in the Fleet

Table 7 demonstrates how the results of the main study can be de-stratified for a hypothetical application population. Of the 1000 vehicles to be sampled, 80 vehicles are expected to be High Evaps as shown at the bottom of Column G. The primary High-Evap-occurrence results of the main study are the nine High Evap fractions shown in Column H. With these nine values and population fractions for any application population (such as the nationwide fleet), the overall fraction of High Evaps in the application population can be calculated as described in Appendix A and as demonstrated by Table 7. For demonstration purposes, this sample calculation uses the California RSD Pilot study population fractions shown in italics in Column D. To estimate the High Evap fraction in an application population, the Column D values must be changed to the actual population fractions for the application fleet in question.

The de-stratification results in Table 7 show that with 1000 vehicles sampled and 80 High Evaps detected, the design can determine the High Evap fraction of an application fleet that is near the actual fleet High Evap fraction. The table shows that the half width of the 95% confidence interval is expected to be 22% of the estimated fraction of High Evaps in the fleet. This precision value of 22% meets the precision target of 25% that the 1000-vehicle design was designed to achieve.

Table 7. De-Stratification for the Main Study Results

A	B	C	D	E	F	G	H	I
Bin Definitions: RSD HC (ppm)			*Application* Vehicle Population	Modified Optimum Stratified Random Design Results			De-Stratifying Calculations	
Bin	**Greater Than**	**Less Than or Equal To**	**Population Fraction in Bin (W_h)**	**Size of Screening Sample Needed to Fill Bin**	**Total Number of Vehicles Sampled in Bin (N_h)**	**Number of High Evaps Found in Bin**	**Fraction of Vehicles in Bin that are High Evaps (p_h)**	**Std Dev of the High Evap Probability for this Bin (σ_h)**
1	-Inf	0	*0.3131*	319	100	0	0.000	0.0000
2	0	90	*0.4281*	234	100	0	0.000	0.0000
3	90	148	*0.0922*	1085	100	0	0.000	0.0000
4	148	245	*0.0711*	1546	110	0	0.000	0.0000
5	245	403	*0.0473*	3422	162	2	0.012	0.1104
6	403	665	*0.0278*	6445	179	9	0.050	0.2185
7	665	1097	*0.0119*	10746	128	20	0.156	0.3631
8	1097	1808	*0.0047*	13709	65	20	0.308	0.4615
9	1808	Inf	*0.0037*	14969	56	29	0.518	0.4997
All			1.0000		1000	80		
							Estimated % High Evaps in Fleet: 0.72%	**Half-Width of the 95% Confidence Interval:** 0.16%
							Half Width of the 95% Confidence Interval as a Percent of the Mean:	22%

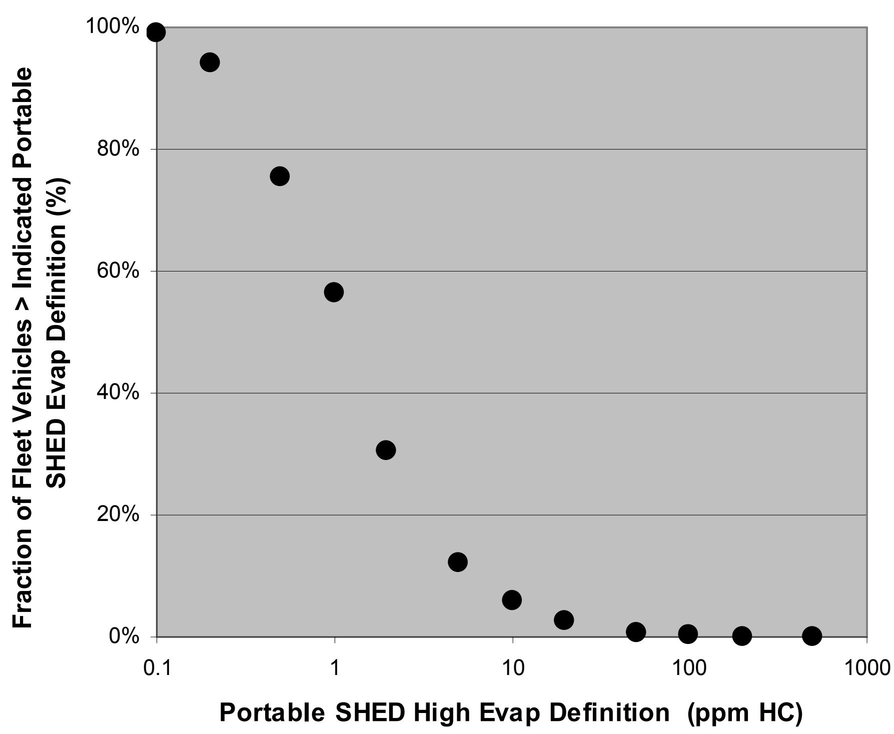
5(b)(ii) Distribution of Fleet Evaporative Emissions

In Figure 3 we interpreted the California RSD Pilot data in terms of the presence of vehicles with high evaporative emissions. In that discussion we talked about High Evaps, even though the definition of "High Evap" is arbitrary. In reality, all vehicles produce evaporative emissions. Most vehicles produce low levels of evaporative emissions, and a few produce high levels.

In the analysis of the actual data collected in this study, we would like to use the portable SHED results to determine the distribution of evaporative emissions for the application fleet as a function of the definition of a High Evap. Then, the fraction of High Evaps in the fleet can be determined from this distribution for any High Evap definition. Alternatively, this distribution can be viewed as the distribution of evaporative emissions of vehicles in the application fleet.

The distribution of portable SHED emissions for an application fleet can be demonstrated by applying the de-stratification factors in Table 6 to the simulated portable SHED results in Figure 5. Imagine a horizontal line in Figure 5 that can move up and down. Points above the line represent High Evaps, and points below the line are non-High Evaps. If the number of High Evaps above the line and within an RSD HC bin are multiplied by the bin's de-stratification factor, the de-stratified High Evap counts for all bins are added together, and this sum is divided by the de-stratified total counts for each RSD HC bin, the result will be the fraction of application fleet vehicles that are above the portable SHED result represented by the horizontal line. Figure 6 shows the result of this calculation for the simulated results in Figure 5 for twelve different portable SHED evaporative emissions levels. For example, the curve shows that when the definition of High Evaps is a portable SHED result of 50 ppmHC or more, 0.75% of the application fleet are High Evaps. Note that the horizontal axis of Figure 6 is a log scale.

Figure 6. Simulated Distribution of Portable SHED Evaporative Emissions for the Application Fleet



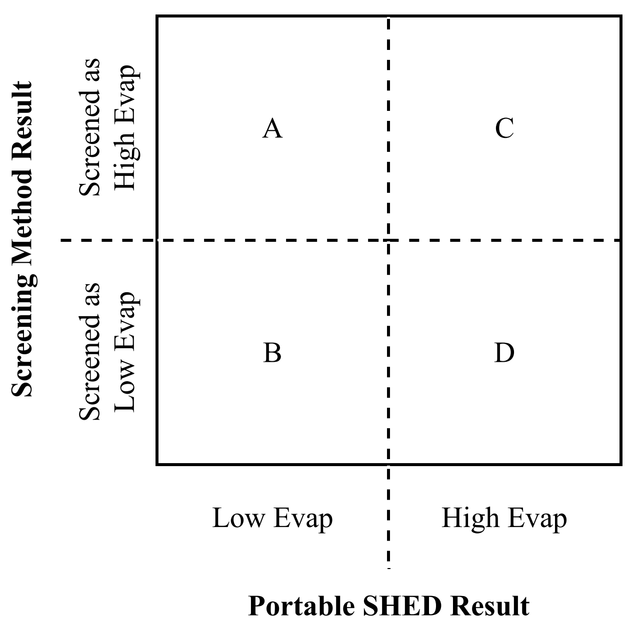
The curve in Figure 6 is also the estimated de-cumulative curve for the distribution of evaporative emissions in the fleet as measured by the maximum HC concentration in the portable SHED.

5(b)(iii) Operating Characteristics of Screening Methods for High Evap Identification

The operating characteristics of each of the screening methods (RSD, modified California method, and infrared camera) will be evaluated against the portable SHED designations of vehicles as High Evaps. The portable SHED will be the standard for designating vehicles as High Evaps in this study. Each vehicle will be designated as either a High Evap or Low Evap based on the measurements made in the portable SHED. That is, in this study the result of the portable SHED will treated as the true evaporative emissions status of each vehicle. On the other hand, each vehicle will be called a High Evap or a Low Evap by each screening method. If all calls by a screening method agree completely with all standard designations by the portable SHED method, the screening method will be regarded as perfect.

Figure 7 demonstrates how the results for each screening method will be evaluated. The portable SHED designation and the screening method result for each vehicle will be plotted. The values for the portable SHED will be continuous numbers. The values for the screening methods will be continuous for RSD, but they will be categorical for the modified California method and the infrared video camera.

Figure 7. Simulated Comparison of Screening Method and Portable SHED Results



The vertical dashed line in Figure 7 separates the portable SHED values into designations of Low Evap and High Evap. The location of this line is arbitrary and will depend on our field experience in the portable SHED while measuring the evaporative emissions of all vehicles tested.

The horizontal dashed line in Figure 7 separates the vehicles into the Screened as High Evap and Screened as Low Evap categories. For the modified California method and the infrared camera method, this line will not be movable since those screening methods will produce categorized, dichotomous results (that is, simply Low Evap or High Evap), and will implicitly incorporate some kind of “cutpoint” even if visual or qualitative. However, since the RSD HC values are continuous, the location of the line depends on the definition (cutpoint) of High Evap that we want to apply to the RSD HC values.

Fur the purposes of determining the operating characteristics of each screening method, it is important to de-stratify the counts of vehicles by using the de-stratification factors in Column G of Table 6. When this is done and the resulting data are plotted in Figure 7, the counts in each of the four quadrants of Figure 7 will be estimates of the relative counts for the application fleet. The counts can be analyzed to determine the operating characteristics of each High Evap screening method. Ideally, all of the data points will be in Quadrants B and C. However, in practice the screening methods will make screening errors. The more accurate methods will make fewer errors.

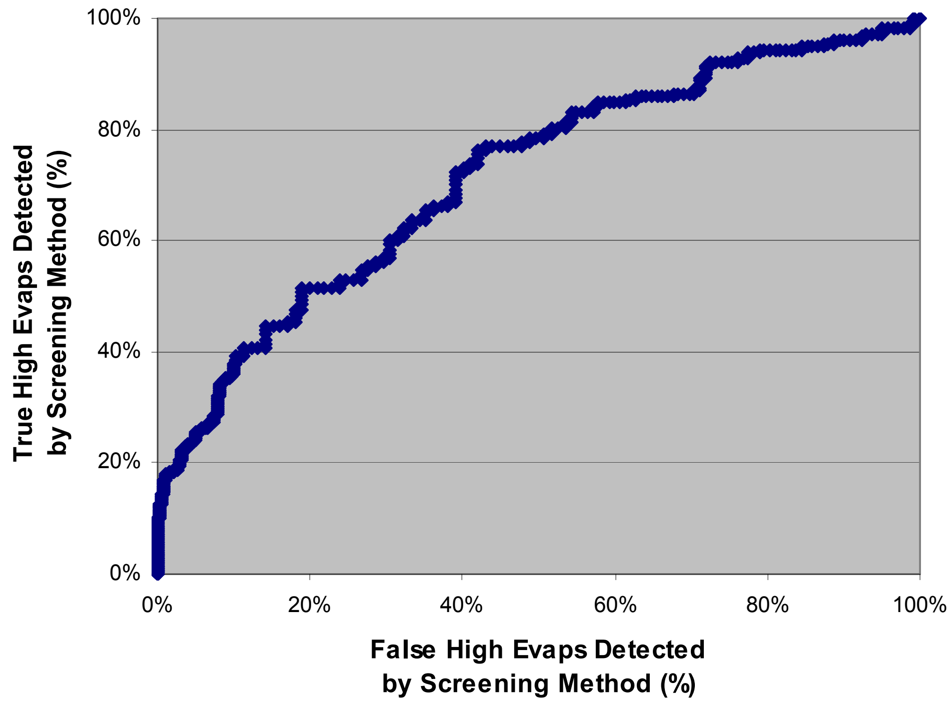
Since the modified California method and the infrared camera method will immediately call vehicles as High Evaps or Low Evaps, their evaluation will straight-fowardly be made by simply examining the counts in the four quadrants. The percent of true High Evaps is defined as C/(C+D), and the percent of false High Evaps is defined as A/(A+B). The screening method with a higher percent of true High Evaps and a lower percent of false High Evaps is the better method.

However, for the RSD method, the analysis is a little more complex since the RSD HC values are continuous. This requires that an RSD HC cutpoint be chosen to separate Screened as High Evaps from Screened as Low Evaps. One way to evaluate for all possible RSD HC cutpoints is to develop the operating characteristic curve of True High Evaps vs. False Low Evaps. The curve is produced by varying the location of the horizontal dashed line in Figure 7. If the percent of true High Evaps and percent of false High Evaps are calculated as the horizontal line is moved from the top to the bottom of Figure 7, the operating curve as simulated in Figure 8 can be produced.

Starting at the lower left in Figure 8, the curve shows that at high RSD HC cutpoints the method may have a high rate of True High Evaps while having a low rate of False High Evaps. Better screening method performance is obtained when the operating characteristic curve passes closer to the upper left corner of the plot.

Operating characteristic curves can also be made for different definitions of High Evap for the portable SHED. Such curves can be used to evaluate the performance characteristics of the RSD method for different High Evap definitions.

Figure 8. Simulated Operating Characteristics of a High Evap Screening Method



REFERENCES

1. D. McClement, "Raw Fuel Leak Survey in I/M Lanes," prepared for American Petroleum Institute and Coordinating Research Council, prepared by Automotive Testing Laboratories, Mesa, Arizona, June 10, 1998.

2. H.M. Haskew, T.F. Liberty, D. McClement, "Fuel Permeation from Automotive Systems," CRC Project No. E-65, prepared for California Air Resources Board and Coordinating Research Council, prepared by Harold Haskew & Associates and Automotive Testing Laboratories, September 2004.

3. D. Amlin, R. Carlisle, S. Kishan, R.F. Klausmeier, H. Haskew, "Evaporative Emissions Impact of Smog Check," prepared for California Bureau of Automotive Repair, prepared by California Bureau of Automotive Repair, Eastern Research Group, de la Torre Klausmeier Consulting, Harold Haskew & Associates, September 15, 2000.

4. L.C. Landman, "Evaporative Emissions of Gross Liquid Leakers in MOBILE6," M6.EVP.009, U.S. Environmental Protection Agency, EPA420-R-01-024, April 2001.

5. A.D. Burnette, S. Kishan, T.H. DeFries, "Evaluation of Remote Sensing for Improving California's Smog Check Program (Version 15 final)," Final Report, ARB-080303, prepared for California Air Resources Board and California Bureau of Automotive Repair, prepared by Eastern Research Group, Austin, Texas, March 3, 2008.

6. R.F. Klausmeier, P. McClintok, "Virginia Remote Sensing Study – Final Report," prepared for Virginia Department of Environmental Quality, prepared by ESP, Tucson, Arizona, February 2003.

7. G.W. Snedecor and W.G. Cochran, Statistical Methods, Sixth Edition, 1967, pp. 216-217.

8. R.O. Gilbert, Statistical Methods for Environmental Pollution Monitoring, Van Nostrand Reinhold, New York, 1987.

Appendix A
Stratified Sampling

The equations pertaining to stratified sampling discussed in this section are presented by Gilbert [8]. The equation for the optimal sample size for a given stratum is:

$$n_h = \frac{n\, W_h\, \sigma_h}{\sum_{h=1}^{L} W_h\, \sigma_h}$$

where

n_h =	the sample size in stratum number h,
n =	the total sample size for all strata,
W_h =	the fraction of the actual population that falls in stratum h,
L =	the number of strata, and
σ_h =	the standard deviation of the distribution from which the individual data values in stratum h are sampled.

This equation follows conceptual guidelines. The number of points taken from a stratum is directly proportional to the fraction of the population comprised of that stratum (the fraction is W_h). Also, the number of points from a stratum is directly proportional to σ_h, which is a measure of the variability in the stratum.

The estimate of the population mean, $\bar{x}_{pop}$, is the weighted mean of the stratum means, $\bar{x}_h$:

$$\bar{x}_{pop} = \sum_{h=1}^{L} W_h\, \bar{x}_h$$

The point here is that the strata are not sampled proportionately to their actual representation in the population. If a simple arithmetic average of the complete stratified sample were calculated, the different strata would be weighted disproportionately to their representation in the population, and a biased average would result. The weighting scheme in the calculation of $\bar{x}_{pop}$ accounts for the nature of the sample and produces an unbiased estimate of the population mean. The formulation here produces the unbiased estimate of the population mean with the minimum error variance, given the total sample size, n. The standard error of the mean is the square root of its error variance. The standard error of this weighted mean estimate is as follows:

$$\sigma_{\bar{x}_{pop}} = \sqrt{\sum_{h=1}^{L} W_h^2 \frac{\sigma_h^2}{n_h}(1 - f_h)}$$

where f_h is the number of data points in stratum h divided by the population size of this stratum.

The factor $(1\text{-}f_h)$ accounts for the finitude of the population in stratum h. If the sample sizes are small compared to the sizes of the strata in the population, this factor can be ignored. The result is somewhat conservative (larger) estimates of the standard errors for the stratified results. The factor $(1\text{-}f_h)$ has been ignored (set to 1) in the calculations presented below, so the standard errors for the stratified analysis are somewhat conservative.

In practice the true standard deviations, σ_h, are not known and are estimated on the basis of historical data that exist before the planned stratified sampling effort. The sample standard deviation, s_h, based on a sample, $x_{h,i}$, i=1 to m, is:

$$s_h = \sqrt{\frac{\sum_{i=1}^{m}(x_{h,i} - \bar{x}_h)^2}{m-1}}$$

where $\bar{x}_h$ is the arithmetic mean.

When the individual data values are dichotomous, for example, 1 for a vehicle with high evap and 0 for a vehicle with low evap, then the standard deviation can also be expressed using the probability p_h that the vehicle has high evap:

$$s_h = \sqrt{p_h q_h}$$

where:

p_h is the probability that a vehicle in stratum h has high evap, and
q_h is the probability that a vehicle in stratum h has low evap ($q_h = 1\text{-}\ p_h$).

Appendix B
Questionnaire

Hello. My name is ________ and I am a contractor for the Environmental Protection Agency. We are conducting a study to understand the evaporative emissions of the vehicle fleet.

We would like to do some testing of your vehicle. The testing of your vehicle would take about 1 hour. Your vehicle was scientifically selected for this study. Your participation is voluntary.

1. Can you tell us approximately how many miles you drive in a given year?
2. Do you park this vehicle inside a garage or outside at night?
3. When was the last time you fueled your vehicle?
4. When was the last time you changed the oil in this vehicle?
5. Have you had any other maintenance performed on the vehicle in the last year?
6. Have you ever had a gasoline smell around your vehicle? If yes, could you describe the circumstance.
7. If yes, have you done anything to fix it?
8. How long have you owned your car?
9. Has the car ever been in an accident to your knowledge?
10. Have you ever replaced the gas cap?

Appendix C
Vehicle Information

1. Vehicle Identification Number
2. Make/Model/ModelYear
3. Odometer reading
4. License plate
5. Evaporative emissions control family
6. Engine displacement
7. Transmission type (manual vs. automatic)
8. Date and Time
9. IM code number
10. A picture of the car (front quarter view)
11. A close-up picture of the license plate
12. A close-up picture of the VIN (could be at the windshield location or some other like the door frame)
13. A close-up picture of the under-hood VECI label

Tracking page to document test completed

- RSD
- Modified CA method
- Infrared Camera
- Portable SHED