APTD-1356

AN EVALUATION OF THE STRATIFIED CHARGE ENGINE (SCE) CONCEPT

U.S. ENVIRONMENTAL PROTECTION AGENCY
Office of Air and Water Programs
Office of Mobile Source Air Pollution Control
utomotive Power Systems Development Division
Ann Arbor, Michigan 48105

APTD-1356

AN EVALUATION OF THE STRATIFIED CHARGE ENGINE (SCE) CONCEPT

Prepared by

L. Bogdan, H. T. McAdams, D. J. Schuring

Cornell Aeronautical Laboratory, Inc.
Buffalo, New York 14221

Contract No. 68-04-0040

EPA Project Officer: J. Dillard Murrell

Prepared for

U. S. ENVIRONMENTAL PROTECTION AGENCY
Office of Air and Water Programs
Office of Mobile Source Air Pollution Control
Advanced Automotive Power Systems Development Division
Ann Arbor, Michigan 48105

January 1972

The APTD (Air Pollution Technical Data) series of reports is issued by the Office of Air Quality Planning and Standards, Office of Air and Water Programs, Environmental Protection Agency, to report technical data of interest to a limited number of readers. Copies of APTD reports are available free of charge to Federal employees, current contractors and grantees, and non-profit organizations - as supplies permit - from the Air Pollution Technical Information Center, Environmental Protection Agency, Research Triangle Park, North Carolina 27711 or may be obtained for a nominal cost, from the National Technical Information Service, 5285 Port Royal Road, Springfield, Virginia 22151.

This report was furnished to the U. S. Environmental Protection Agency by Cornell Aeronautical Laboratory, Inc. in fulfillment of Contract No. 68-04-0040 and has been reviewed and approved for publication by the Environmental Protection Agency. Approval does not signify that the contents necessarily reflect the views and policies of the agency. The material presented in this report may be based on an extrapolation of the "State-of-the-art." Each assumption must be carefully analyzed by the reader to assure that it is acceptable for his purpose. Results and conclusions should be viewed correspondingly. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Publication No. APTD-1356

FOREWORD

This report was prepared by Cornell Aeronautical Laboratory, Inc. (CAL), Buffalo, New York, under Environmental Protection Agency (EPA) Contract No. 68-04-0040, Modification No. 4. The work was administered under the direction of the Office of Air Programs, Division of Advanced Automotive Power Systems Development, Mr. J.D. Murrell, Project Officer.

This is the Phase II interim technical report that describes and summarizes the results of studies conducted during the period from October 1971 through December 1971 on the engineering and emissions performance aspects of the stratified-charge engine concept.

The effort was performed jointly by the Vehicle Research Department and the Systems Research Department of CAL.

ABSTRACT

A descriptive account is presented of preliminary studies aimed at establishing a rational basis for the evaluation of the stratified charge engine as a low-emission power plant for light duty vehicle applications. A description of the salient engineering features of the stratified charge engine is included together with statistical analyses of the limited exhaust emissions data that are available on vehicles equipped with hand-built experimental engines. For purposes of augmenting this limited data base and for purposes of comparison and analogy, statistical analyses of pertinent emissions data obtained for conventional-type automobile engines are also presented.

To provide essential but missing data, the rationale for and the details of an experiment design for presently available engine/vehicle units is presented. New hardware requirements are also identified.

TABLE OF CONTENTS

TABLE OF CONTENTS (cont'd)

I INTRODUCTION

This report summarizes the results of preliminary studies conducted for the Division of Advanced Automotive Power Systems Development, Environmental Protection Agency, and directed to the ultimate objective of establishing the technological feasibility of the stratified charge engine (SCE) concept meeting the 1976 Federal exhaust emission standards in light duty vehicle applications. In evaluating technological feasibility, such important factors of the engine/vehicle composite as performance, drive-ability, manufacturability, serviceability, and durability will need to be considered in addition to the emissions aspects.

The SCE is a hybrid power plant that combines certain key features of spark-ignition and compression-ignition engines in such a way that some of the more desirable operational characteristics inherent in each design philosophy are realized. Fuel injection with charge stratification permits operation at lean air-fuel ratios (producing significant fuel economies) and the use of high compression ratios (without correspondingly increased fuel octane requirements) facilitates more efficient power output over a limited speed range than the homogeneous charge (carbureted) engine. In contrast with the diesel, positive ignition produces smooth combustion, permitting a compact, lightweight engine design that combines low cost with good cold-start characteristics.

These features of the SCE have promising logistics potential for military purposes and the U.S. Army Tank Automotive Command (USATACOM) has been sponsoring the development, both by Ford and Texaco, of an SCE replacement for the standard 4-cylinder military jeep engine. When tests on the early, hand-built stratified charge engines showed that low exhaust emissions were inherent in this design, the EPA contributed financial support to the USATACOM program to help in the further development of this engine concept. While the EPA and the USATACOM share many common objectives in the engine development program, there are some divergent goals The EPA is concerned with the ultimate feasibility of mass producing more powerful SCE versions (six or eight cylinder engines, for example), their compatibility with the requirements of passenger vehicle applications, and their durability and exhaust emissions performance in terms of compliance with the 1976 standards. Accordingly, the EPA is formulating an evaluation program that embodies its particular objectives. The study summarized in this report represents an initial phase. While the objectives of this present study were broadly drawn, nevertheless, the primary emphasis has been on the emissions performance of the engine.

Specifically, the study has comprised a literature survey designed to consolidate information on critical design factors, components, and parameters affecting SCE performance and exhaust emissions together with indications where trade-offs and compromises are feasible. Included also are statistical analyses of available emissions data on the 4-cylinder SCE to extract such information as: repeatability of test results, differences between sample bags (of the 3-bag CVS test procedure), the statistical significance of engine adjustments on engine emissions, and possible effects of ambient test conditions. Because of the extremely limited quantity of data on the SCE, statistical analyses were conducted on selected emissions data for conventional homogeneous charge, spark-ignited, internal combustion engines to evaluate such factors as reproducibility of test data within a model and within model types, deterioration of emissions with mileage, ambient effects on emissions and, by analogy, to attempt to relate these data to the SCE.

With these efforts delineating the type of important data and information that were presently unavailable, the considerations and rationale were specified for an experiment on existing hardware to provide some of these data. Existing hardware in this case refers to several 4-cylinder SCE installed in 4-speed manual shift transmission, 4-wheel drive, military jeeps and one 2-wheel drive Postal Service van with 2-speed automatic transmission. In addition, new hardware requirements for assessing emissions performance of the SCE were defined from the standpoint of how many units must be built and tested to have "reasonable assurance" that the developmental SCE will perform satisfactorily.

II SUMMARY

A survey of the literature related to the Ford (FCP) and Texaco (TCP) stratified-charge engine concepts indicates there are operational options possible by adjusting fuel injection and spark ignition timing. The timing schedule can be selected to avoid preignition and detonation, leading to multifuel capability but poor full power output and smoke problems. An alternative schedule achieves excellent full power output at the expense of susceptibility to knock and preignition. Critical components, both from the standpoint of durability and feasibility of mass production, are the fuel injection pump and nozzles. Due to the use of catalysts, lead-free fuel is required by stratified-charge engines.

The FCP engine, installed in an Army jeep and equipped with emission control devices, has demonstrated operation at exhaust emission levels which at low mileage are within the 1976 goals. This performance has been achieved at the cost of a 30% power loss and resultant poor drive-ability. Analysis of limited emissions data has yielded a value of approximately 15% for the coefficient of variation associated with the "within" engine variability for each of the three exhaust gas components. In the absence of any data, a similar value (15%) has been assumed for the coefficient of variability associated with the "between" engine variability.

To provide a broader data base from which to draw inferences concerning the stratified-charge situation, emissions data from tests on conventional engines were also analyzed. These data showed that:

(1) Significant differences can exist between two different makes of vehicles although good engine-to-engine repeatability can be achieved.

(2) Conventional vehicles (with no emission controls) failed to show any deterioration of emissions with mileage.

(3) A wide variation in test error was noted in results obtained at different laboratories and with different vehicles. Since these variations may mask other effects under investigation, it suggests that the emissions measuring procedure needs to be improved.

Experiment designs appropriate to testing the limited number of available stratified-charge vehicles have been delineated. The tests involve a study of durability and the effect of parametric variations on emissions. A limited number of additional vehicles are recommended to establish the engine-to-engine variability which at present is unknown. Assuming a 15% coefficient of variation is appropriate to both the test-to-test and engine-to-engine variability of the SCE, a total of ten engine/vehicle units would suffice to define mean emission data with a 95% assurance of being within $\pm$ 10% of the expected true population mean.

Following tests of the limited quantity of hand-built engines available and on order, it is recommended that a like quantity of soft-tooled engines be acquired for intensive evaluation.

III STUDY RESULTS

A. General

A somewhat unconventional structure is employed in this report wherein the discussion text consists of brief, summary-type accounts of the various technical studies conducted during the program. Detail descriptions of the individual studies, together with numerical results where appropriate, are included as appendices.

Use of this report form has been occasioned by the natural evolution of the program phase. Documentation of the developmental studies and test results obtained with the prototype SCE has been meager to date. Accordingly, it was necessary, with the help of the EPA, to identify, locate, and obtain potentially useful items of information and test data. As these separate inputs were received, analyses and studies were immediately made and the results documented in a form that would constitute an appendix for this report.

B. Survey of Available Literature on the SCE

A literature survey was made with the intent of consolidating technical information on the SCE related to critical design factors, components, and parameters insofar as they affect operational performance and exhaust emissions. An additional function was the identification of compromise tradeoff possibilities and their consequences, together with an indication of the known problem areas. While the stratified-charge concept dates back several decades and its history includes the developmental efforts of a number of individuals and organizations, only two specific implementations have been considered here. These include the Texaco Combustion Process (TCP) and the Ford Combustion Process (FCP) which are currently in competition for selection as the SCE replacement for the military L-141 jeep engine. Physically and conceptually, the two approaches are very similar with the differences reflecting the differences in the stated design objectives. While many of the objectives such as good fuel economy, for example, were similar, the principal difference lay in the fact that the TCP focused on multi-fuel operational capability while the FCP goals dwelt more on low emissions, performance and smoothness of engine operation.

Proper stratification of the fuel-air mixture, so as to achieve a rich, easily ignitable mixture at the plug electrodes and leaner but still combustible mixtures at the periphery of the charge, is the key factor in the successful realization of a useful SCE. The SCE concept requires in its implementation the proper coordination of three basic parameters: air swirl, fuel injection, and positive spark ignition.

Air swirl rate is a vital parameter in the stratification process and affects both the efficiency and the duration of combustion. Such geometrical devices as shaped inlet ports and shrouded intake valves are employed to create an initial low-rate swirl in the cylinder. A combustion-cup-in-piston-crown configuration achieves the required high swirl rate near the end of the compression stroke.

Experiments with various cup geometries have shown the largest effects to be on fuel economy and combustion harshness. Cup designs incorporating sharp edges or ridges have proven most satisfactory and the speculation exists that these surface discontinuities promote turbulence and mixing which in turn creates smooth combustion. If this supposition is correct, the accumulation of combustion chamber deposits with usage could serve to round-off these surface discontinuities and produce unfavorable engine operation.

The military development objective of maximum parts commonality with the standard L-141 jeep engine necessitates the use of an "oversquare" bore/stroke ratio which is unfavorable to swirl creation. For this reason, as well as others discussed in Appendix A, it would be desirable to design SCE prototypes optimized for best engine performance, unconstrained to utilize design parameters and components of an existing engine designed for homogeneous-charge combustion.

Another critical factor in the operation of the SCE is the fuel injection process which must reproducibly create the proper fuel spray pattern. The fuel injector is the critical component here and it is susceptible to "coking" and fouling with attendant variations in the fuel spray pattern. Durability data are not yet available to determine the severity of these potential problems.

Basically the SCE may be operated in an unthrottled condition with the engine power determined by the time duration of the fuel injection. While the earlier versions of both the TCP and FCP engines were operated in this mode, poor idle performance has resulted in both systems incorporating some air-intake throttling over the entire speed range of the engine

In the FCP engine a more or less conventional spark ignition system is employed. A spark plug with long electrodes was used in early development engines to locate the ignition point within the fuel-rich core of the swirling charge. In later (PROCO) versions of the engine employing air throttling and exhaust gas recirculation, this ignition point appeared more fuel-rich than desirable for flame initiation. A short-electrode plug geometry therefore was developed that resulted in improved operation and lower HC and CO emission levels. Severe spark-gap erosion has been encountered because of the use of a high-energy ignition system needed to preclude misfire at part-power engine operation

The TCP version uses a special transistorized ignition system with a number of unique characteristics among which is the ability to generate a constant-voltage, controlled-duration spark. This feature makes the ignition/injection timing less critical.

Control of combustion in the SCE depends on the relative timing of spark ignition and fuel injection with respect to crank angle during the compression stroke. The TCP and the FCP versions of the SCE differ most in this respect and illustrate the consequences of each philosophy. In the Ford system independent advance-retard timing is employed for ignition and fuel injection as a function of engine speed and load. This scheme achieves excellent air utilization at full load operation and hence realization of full power output without the smoke problem normally associated with heterogeneous charge engines under these circumstances. The price paid for this advantage is that the engine is susceptible to preignition and detonation. Hence useful compression ratios are limited and fuel octane requirements become more critical. To achieve and maintain the critical ignition/injection timing schedule, Ford has developed a unitized distributor and fuel injection pump for this purpose.

The Texaco system appears* to be timed so that the first increment of fuel injected is <u>always</u> ignited. Thus the possibility of preignition and detonation is precluded quite independently of the fuel octane rating and a

---

* The literature is not explicit on this point.

multi-fuel capability for the engine is achieved. On the other hand, air utilization is poor at high load conditions so that maximum output is exhaust-smoke limited.

Both the TCP and the FCP versions, as 4-cylinder, hand-built prototypes installed in jeeps, have successfully demonstrated attainment of their objectives providing: significant fuel economies, multi-fuel capability (TCP only), cold start capability, immediate driveaway after cold start without warmup, excellent throttle response, and torque/power output equivalent to the L-141 engine. Exhaust emissions (HC and CO) are smaller than for a conventional engine since such causes of emissions as wall-wetting of intake manifold by fuel, flame quench at combustion chamber walls, and deficiency of air are either eliminated or minimized. Oxides of nitrogen are about at the same levels for the 4-cylinder SCE and the standard L-141 unit.

EPA tests on a 4-cylinder FCP installed in a military jeep have produced low-mileage emission levels lower than the Federal 1976 emission standards when the engine is equipped with exhaust gas recirculation (EGR) for control of NO_x and an exhaust system which includes a thermal reactor and a catalyst for control of HC and CO. Engine adjustments to achieve this level of emissions were such that approximately a 30% power loss was experienced with occasional misfire at high speed. The exhaust gases of the SCE have an unpleasant odor which can also irritate the eyes and nose. The source of this odor is believed to be associated with aldehydes and some corrective and control measures have been proposed. No demonstrable success has yet been achieved.

In addition to studies with single-cylinder and four-cylinder prototype SCE engines, Ford has also converted some large V-8 engines to a stratified-charge configuration. Little detailed information is available since this latter effort constitutes company-funded research. These engines appeared to operate adequately but with certain demonstrated deficiencies which would not classify them as acceptable in luxury-car applications. These deficiencies are presumably correctable.

Recent discussions (January 1972) with Ford personnel active in the FCP program indicate that scaling or extrapolation of performance or emissions data from a single-cylinder to a four-cylinder SCE or a four-cylinder to an eight-cylinder SCE is complex and unreliable. One factor is that each engine size has to be operated with its own unique injection pump. Attempts to use one common pump on different size engines have not been successful. Another factor concerns the pressure pulsations in the intake manifold which cause interaction between cylinders adversely affecting swirl formation and air induction. With the lack of adequate theory in engine development, all approaches reduce to "cut-and-try" type operations.

If the SCE is to fill the role of a satisfactory passenger vehicle engine and simultaneously meet the emission goals for 1976, two critical questions must be resolved: (1) the feasibility of production of the SCE in mass quantities to the required tolerances, and (2) the feasibility of the engine/emissions control unit maintaining tolerably-low emissions over a 5-year/50,000-mile endurance period. Critical production items are the injection pump and the injection nozzles. Durability-related issues that are unique to the SCE include injection pump deterioration and injector fouling resulting in inability to operate with consistent reproducibility. Of lesser import, but still of concern, is the matter of exhaust odor.

A detailed review of the literature survey study on stratified-charge engines is included in Appendix A - Summary of a Literature Survey of the Stratified-Charge Engine Concept.

C. Statistical Analysis of Selected Test Data

1. Stratified-Charge Engines

Evaluation of the feasibility of mass production of SCE-equipped vehicles meeting the 1976 emission goals requires sufficient data on a representative sample of these units so that the population distribution (i. e., the population mean together with some measure of the dispersion relative to the mean) of the ultimate production fleet can be predicted with a reasonable level of assurance. At the present stage of the SCE development, neither such data nor such a sample of vehicles exists. The rationale and the design details of structured experiments to obtain such data are discussed in subsequent portions of this report. Critical to the design of such an experiment is a knowledge of the repeatability of the emissions performance of a vehicle when subjected to replicative tests (i. e., the with-in vehicle/engine variance) and the reproducibility of emissions among different units of the same models of the vehicle/engine (i. e., the between vehicle/engine variance). The strategy of the experiment depends on the relative levels of these two variances. Should the with-in vehicle/engine variance be relatively the larger, then the methodology would dictate many replicative tests on relatively few test units. If the converse situation prevails, then fewer tests would be performed on any one vehicle/engine unit but more units would be tested.

When tests are replicated on a single vehicle/engine unit, a certain scatter or dispersion occurs in the resultant emissions measurements. The sources of this random scatter reside in the inability of the engine to repeat itself (the with-in engine variance) plus irregularities in instrumentation, deviations from prescribed test procedures, operator variability, variations in ambient conditions, etc. All these sources combine to produce random scatter in the data and the sought-for within-engine variance is confounded with the variances from other sources. When tests on different vehicle/engine units are made, the measured results are further confounded by the between-engine variance. If necessary, the variance components attributable to the individual sources may be separable by use of the analysis of variance methodology.

A limited quantity of emissions test data were available on an FCP engine installed in a military jeep. These data, consisting of 14 tests conducted by EPA, Ypsilanti, and 9 tests conducted by Ford, were supplied by the EPA. They consisted of emissions (grams/mile) for HC, CO, and NO_x as measured by the 3-bag, CVS-CH procedures using the LA-4 driving schedule. During the course of these tests several changes were made to the EGR linkage, fuel enrichment at WOT, and deceleration fuel cut-off. A change of catalyst was also made early in the test series.

Using statistical techniques, the entire data set was tested for homogeneity (i.e., whether the data belonged to the same population distribution). The data identified as homogeneous was then pooled and the means and standard deviations for each pollutant were computed. From these results, the coefficient of variation (defined as the ratio of the standard deviation to the mean, expressed as a percentage) was computed. For all three exhaust gas components the coefficient of variation was essentially the same, 16% 17%.

It will be appreciated that the standard deviations computed include not only engine variations but also those due to test apparatus, test procedures, etc. Whether these other sources contribute significantly can not be ascertained at this time.

Details concerning the test data, calculations, and other analyses aimed at evaluating effects of engine adjustments on emissions performance, as well as a comparison of these data with 1976 standards, are all included in Appendix B, Analysis of Emissions Test Data on Low-Emission FCP Engine Installed in M-151 Vehicle.

The emissions data could have been influenced by ambient factors such as atmospheric pressure, specific humidity, and air temperature, as well as such imposed variables as dynamometer inertial load and horsepower. To estimate the effects of these variables on emissions, regression analysis was applied to the data from the set of 14 tests conducted by the EPA. A linear regression equation was fitted to these data using the least squares method and the statistical significance was tested by the analysis of variance. Results of this analysis indicate that the data from these 14 tests do not establish any linear dependence of emissions on atmospheric pressure, specific humidity, temperature, load and horsepower. Since the combined effects of the five variables fail to attain a statistically significant level, it can also be inferred that any single variable would similarly fail to demonstrate a significant effect on emissions. Thus the coefficient of variation calculated from these data may be expected to be free of any systematic influences due to these five enumerated variables. Appendix C contains a detailed account of the calculations conducted in performing the regression analysis and the analysis of variance.

It is important to be aware of the fact that these emissions tests conducted on the SCE were not structured to establish the effects of these variables on emissions. That is, no deliberate attempt was made to vary any of the five parameters over a range of values. Consequently, only a naturally occurring small variation in the ambient environmental conditions was experienced and the chassis dynamometer settings also encompassed a limited range of adjustments. Consequently, the inferences drawn here only apply to the particular set of circumstances applicable to these test results. To prove the point, it is an established fact that specific humidity does affect NO_x emissions and a correctional procedure is included for these effects in the emissions test calculations applicable to 1976 standards (the raw data used in this analysis were uncorrected data).

2. Conventionally-Powered Vehicles

The quantity of emissions data available on stratified-charge engines is very meager and all of the engines are different from each other in some respect. Consequently, this limited data base does not permit inferences to be drawn concerning variability in emissions attributable to differences existing among identical engines of the same make (i.e., a measure of product variability). To help in augmenting this limited data base, selected emissions data that were available on current (and recent) models of conventionally-powered vehicles were statistically analyzed. It was felt that this type of information would be useful in the planning of test programs for the SCE.

a. Test Error

This one part of this effort focused on relating the magnitude of the test error to emission differences ascribable to product variation and to the other various sources of variation involved. "Test error" is here described as the measured fluctuations obtained by repeated tests performed on the same vehicle by the same equipment under controlled test conditions. The data encompassed a wide variety of vehicles: new and with accumulated mileage, with and without emission control devices, and even conventionally-powered military jeeps. For comparison purposes, data on the SCE-equipped jeeps were also included.

The emissions data were all found to be non-normally distributed but rather were represented by skewed distributions. To a good approximation, all of the data were found to follow a log-normal distribution. These data can be presented most easily by plotting the relative cumulative frequency of occurrence on probability paper whose abscissa scale is logarithmically graduated. A log-normal distribution will plot as a straight line on such graph paper with the slope of the line indicative of the dispersion of the data. Data from the various sources, for each of the three components of the exhaust gases (HC, CO, and NO_x) could be represented by such graphical techniques. Depending on the data source, the graphical data could include not only test error but other sources such as product variability, deterioration of performance with usage, variations associated with different makes of engines, and the like. It would thus be logical to expect that the larger the number of sources of variability operable within a given data set, the larger would be the dispersion in the data. In graphical form, this would be equivalent to expecting the linear plot to have a shallower slope.

To test the reality of this supposition, all of the graphical data were standardized so as to be presentable on a single sheet of graph paper with all plots sharing a common intersection (with coordinates corresponding to an emission of 1 gram/mile and a cumulative frequency of 50%) and preserving their original slopes. While the data in general supported expectations, there were a significant number of exceptions. Data sets with few known sources of variability often exhibited a larger dispersion than data sets with many more sources of variability involved. It is suspected, therefore, that the lack of consistency is associated with the test error, which is present in all measurements, but apparently is subject to sufficient variations with vehicle make, vehicle service, measurement techniques and procedures so as to obscure other active sources of variation. The conclusion is that the test error can be highly variable but evidence also exists to show that it can be made small. Data taken on the SCE showed a relatively small test error but even smaller levels of test error were observed. Hence, no basis exists for assuming that the stratified-charge engines would show a significantly different test error than an ordinary, carbureted engine.

A discussion of this analysis together with tabulations of the numerical data employed and the corresponding graphical presentations are included as Appendix D, entitled Analysis of Emission Test Data of Various Groups of Automotive Vehicles.

b. Parametric Effects and Test Procedures

A separate and distinct set of emissions data on fifty-four 1970 model fleet cars was also received from the EPA. This data set included cars of various models from different manufacturers and had varying amounts of accumulated mileage. Tests had been conducted using the CVS cold-start and CVS hot-start procedures. A four-bag collection system was used: hot start transient, hot start stabilized, cold start transient, and cold start stabilized. These data permitted the following questions to be addressed: does mileage affect emissions, is there a significant difference between emissions from a cold start versus a hot start, and how are the CVS-C and CVS-CH test procedures numerically related? This information, at least inferentially, was considered as instructive in planning an evaluation program for the SCE.

An analysis of variance failed to establish any correlation between emissions data and mileage which for these vehicles spanned the range from 4,000 to 35,000 miles. Similar analyses also failed to show any trends of emissions with number of cylinders, displacement, and vehicle weight which were the other parameters given (because of commitments to hold some aspects of the data as confidential, EPA was not at liberty to identify the vehicles except by range of engine displacement, number of cylinders, and approximate weight). These analytical procedures are presented in part 2 of Appendix C.

Analysis of the individually-bagged emissions established the fact that the HC and CO contents of the first (cold start) bag are significantly higher than the other three. Differences among the second (cold stabilized), the third (hot transient), and the fourth (hot stabilized) were found to be negligibly small (for HC and CO). Compared to the average of the subsequent bags, the HC mean of the cold start bag is approximately 1.6 times larger; similarly the CO mean was 3.4 times larger. In the case of NO_x, the transient bag emissions always exceeded those of the stable bag, irrespective of whether the start was hot or cold.

An investigation of the effect of bag weighting on the computed mass emission level showed that the CVS-CH procedure results in an HC level that is 12% lower than the CVS-C method. The CO level is lower by 29% while NO_x levels are virtually unaffected. Calculations leading to these latter results are given in Appendix E, Emissions Analysis of 54 Fleet-Type Passenger Vehicles.

c. Product Variability

A report published by Automotive Research Associates (ARA) of San Antonio, Texas, contains results of emissions tests on five 1969 model Fords and five 1969 model Chevrolets conducted at 10,000-mile intervals over a total range of 100,000 miles. Since these data failed to indicate a systematic trend of emissions with vehicle mileage, differences observed in the successive (10,000-mile tests) can be attributed to random causes due to testing errors and actual changes in vehicle performance from test to test (the "with-in" vehicle variance). By comparing the "between" vehicle variance and the "with-in" vehicle variance, an estimate can be made of the effect that individual differences among vehicles might contribute to emissions assessment. Since such differences might be expected, in part at least, to reflect effects of manufacturing tolerances, the separation of "between" and "with-in" variability could permit speculation on the effects of manufacturing tolerances on SCE emissions.

Using analysis of variance techniques, the emissions data for Ford and Chevrolet vehicles was used to compute two sets of F-ratios. The one set relates to the question of whether the "between" variance for each make was significant at the 0.05 level. The other set relates to the question of whether significant differences exist between the variability of the two makes of vehicles. These differences are of two types: one type concerns the ability of the vehicle to reproduce its performance on successive tests while the other type concerns reproducibility in the vehicle-to-vehicle sense. This latter consideration could be indicative that the manufacturing variability (tolerances) is greater for one manufacturer than the other.

The analytical results indicate that the Chevrolets tested were more variable than were the Fords, and appreciable emissions variations can exist among vehicles of the same make. Between-vehicle variability thus is not always negligibly small. The fact that one of the vehicle makes in these tests exhibited less variability than the other is indicative that the achievement of good engine-to-engine reproducibility is feasible but not always encountered in practice. Whether this observation can be extrapolated to the situation of the stratified-charge engine is not presently known. An account of the details of the analysis of the ARA data, together with tabulated results, is given in Appendix F, Exhaust Emissions for Chevrolet and Ford Automobiles.

D. Statistical Modeling Considerations

Exhaust emissions from stratified-charge engines depend on a number of variables. Some of these variables can be controlled and their effects on emissions determined by systematic experimentation. Other variables are of such a nature that their control is impractical. The effect of these variables on emissions can best be assessed in statistical terms.

In order to put into perspective the various influences which can affect emissions, it is desirable to have a conceptual model which delineates how the various contributions to variability combine to produce the total variability which is observed when many engines are tested under many different environments and use conditions. Moreover, it is desirable to have a rationale whereby the results from a limited number of tests can be used as a bssis for inferring what might be observed if a larger number of tests were conducted. Finally, such a model is essential to the structuring of additional tests or experiments to fill information gaps in available data.

In Appendix H, Experiment Design Considerations for Emissions Testing of Stratified Charge Engines, the influences of the factors which affect emissions are treated as either fixed effects or random effects. It is assumed that the several contributions to variability can be considered as if they combine in an additive manner. A simple expression of this additivity is the equation

$$x_{ij} = \mu + a_j + \epsilon_{ij}$$

One of the uses of this equation is in connection with replicate or repeat tests on two or more vehicles. If j indexes the vehicles and i indexes the test on each vehicle, all test results can be referenced to their mean value μ , which can be regarded as a contribution common to all tests. The j^{th} vehicle has a contribution a_j peculiar to it, and, within that vehicle, the repeated tests make a contribution ϵ_{ij} peculiar to the i^{th} test. A clear implication of the model is that no amount of replicate testing on a single vehicle can provide information on the vehicle-to-vehicle variability represented by the term a_j . A full discussion of an approach to the analytical treatment of sources of variability is given in Appendix H. The implications of these sources of variability for compliance with emissions standards requires statistical consideration.

Statistical treatment of the variability in test results from stratified-charge engines requires the assumption of a statistical distribution which is believed to be applicable to the data. Classically, it is assumed that for a random-effects model the contributions a_j and ϵ_{ij} are governed by the normal probability distribution. Under this assumption, the usual procedures of analysis of variance can be employed and their findings evaluated according to standard statistics.

As shown in Appendix D, Analysis of Emission Test Data of Various Groups of Automotive Vehicles, emissions data tend to follow not the normal distribution but the log-normal distribution. Reasons for this behavior are explored in Appendix G, The Log-Normal Distribution as a Statistical Model for Exhaust Emissions. In this appendix it is argued that emissions data can be expected to tend toward the log-normal form if several sources of variability combine multiplicatively rather than additively. The argument is based on the fact that if one considers not the actual contributions to variability but the logarithms of these contributions, then multiplication of contributions is converted to addition of their logarithms.

Then, by virtue of the Central Limit Theorem of mathematical statistics, it might be argued that the combined logarithms would tend to be normally distributed. Moreover, the additive model for combining sources of variability would now be applicable, provided that it is applied to the logarithms of the emission data.

Considering emissions in terms of a normal distribution in logarithm space is completely equivalent to considering these emissions in terms of a log-normal distribution in untransformed (antilog) space. Indeed, there are straightforward relationships between the parameters of the two distributions. For example, the mean in logarithm space is analogous to the median in antilog space, and the standard deviation in log space is analogous to the "ratio standard deviation" in antilog space. For example, if one wishes to compute (approximately) the 95th percentile of a set of emissions measurements, the following are equivalent procedures:

Log Space:	Mean + Standard deviation + Standard Deviation = Mean + 2 Standard Deviation
Antilog Space:	Median x Ratio Standard Deviation x Ratio Standard Deviation = Median x (Ratio Standard Deviation)^2

A further relationship, and one which provides a very useful approximation, is the relationship between the ratio standard deviation and the quantity called the coefficient of variation. It will be recalled that the coefficient of variation is the ratio of the standard deviation to the mean for a set of data. It is a useful quantity in the assessment of emissions because the standard deviation tends to be proportional to the mean. It is shown in Appendix G that, if the coefficient of variation is small (say $\leq$ 0. 25), then

Ratio standard deviation $\approx$ 1 + coefficient of variation.

If the coefficient of variation (or relative standard deviation) is used in the making of statistical inferences, on the assumption of a normal distribution, therefore, these inferences will often be substantially the same as those evolving from the use of the log-normal assumption.

E. Generation of New Data

Emission data for conventional internal combustion engines have been examined with a view toward drawing inferences which may, to at least a limited extent, apply to stratified-charge engines. It must be admitted, however, that such data, at best, aids primarily in understanding the interrelations of the several sources of variability. Inasmuch as the stratified-charge concept has no real precedent in conventional vehicles, it is essential to obtain a data base drawn from tests on actual stratified-charge hardware. Only a few experimental units incorporating the stratified-charge concept have been built to date and, for the most part, emissions tests have been conducted by EPA on only one unit. In order to supplement this very limited data base in an effective way, it is essential to design the experimental tests in accordance with known data requirements and so as to reflect postulated statistical models. In view of the need for better assessment of differences among individual engines or engine-vehicle combinations, an additional consideration is the question of new hardware requirements. The structuring of tests for existing hardware is discussed in the next section; following that is a discussion of new hardware requirements to provide an adequate data base.

1. Experimental Design for Existing Hardware

A full discussion of a proposed testing program for existing stratified-charge engines is presented in Appendix H. The testing program is based on the availability of one 4-wheel drive jeep equipped with Ford Combustion Process (FCP) engine, one 4-wheel drive jeep equipped with Texaco Combustion Process (TCP) engine, and one 2-wheel drive postal van, equipped with FCP engine and automatic transmission. A few benchmark tests of a standard (conventional engine) jeep are included as reference.

With such limitations in available hardware, it is important to consider carefully which questions can be legitimately attacked by immediate tests and which questions had best be deferred until more hardware is available. For example, with only a few vehicles available, it is virtually impossible to obtain much information pertaining to vehicle-to-vehicle variations. One can, however, consider tests aimed at evaluating mileage degradation effects on emissions. Tests aimed at observing the effects of parametric changes in operating conditions can be considered, but it should be remembered that design concepts are still in a state of flux and that conclusions drawn from such tests should be considered tentative.

On the basis of such considerations, a test array was devised (see Appendix H). The parametric variations incorporated in the testing program are directed toward at least preliminary study of the effects of catalyst, inertial weight, EGR, and mileage on emissions. The schedule set forth includes both complete, cold-start tests and tests employing only the hot-start portion of the test cycle. In this way, it is believed that economy of time and effort can be realized without seriously jeopardizing the validity of the tests. An important feature of the proposed test program is a set of tests in which the engine is "rejuvenated" after 50,000 miles of operation. This rejuvenation would consist of such simple remedial measures as retuning and replacing sparkplugs. The intent of the tests on the rejuvenated engine is to see how much of the observed degradation in emissions characteristics is reversible.

2. Definition of New Hardware Requirements

The question of how much data is required in order to make an assertion at a specified level of confidence is a fundamental consideration in the definition of new hardware requirements. The question has two aspects: the statistical repeatability of emission measurements and the permissible width of a confidence interval for the expected or mean value of the emissions measurements. An approach to this problem is presented in Appendix I, New Hardware Requirements for Assessing Exhaust Emissions Performance of the Stratified-Charge Engine.

The statistical approach of Appendix I is based on the assumption that, for a given category of emission tests, the standard deviation of emission measurements tends to be proportional to the mean value of the measurements for that particular category. Generally speaking, the magnitude of the variation will depend on the opportunity afforded for various sources of variability to come into play. For example, replicate tests performed on the same engine would be expected to exhibit less variability than an equal number of tests performed on different engines. Data supporting this observation is provided by Appendix D, Analysis of Emission Test Data of Various Groups of Automotive Vehicles. For a particular set of circumstances, therefore, it will be assumed that the ratio of standard deviation to the mean is approximately constant and that this ratio, called the coefficient of variation, is an important quantity influencing the number of vehicles which should be procured for further testing.

In the absence of definitive measures of the coefficient of variation for realistic sets of data for stratified-charge engines, one can explore parametrically various assumptions about its value. In Appendix I, a table is provided which tabulates the width of a 95% confidence band for various coefficients of variation and for various numbers of tests. For example, if it is assumed that test-to-test variability for a particular engine yields a 15% coefficient of variation, then nine replicate tests would be required to narrow the half-width of the 95% confidence interval to approximately 10% of the mean value. An alternative statement might be that nine replications are required to define an emission value which has 95% assurance of being within ± 10% of the "true" or expected value.

The same type of argument can be applied to variations from one engine to another. Let it be assumed that for a particular engine a sufficient number of replicate tests can be conducted to narrow the confidence interval to negligible proportions. Nevertheless, if a population of vehicles varies considerably from one vehicle to another, many vehicles would need to be tested in order to estimate the expected value for the population to within a specified confidence interval. Again, if vehicle-to-vehicle variation is such that the coefficient of variation is 15%, then nine vehicles would need to be procured in order to estimate the expected value for all the vehicles to within ± 10% of its true value. It is on the basis of this argument that it is recommended that no fewer than ten vehicles of each type (FCP and TCP) be acquired to develop an emissions data base.

Needless to say, there is no assurance that stratified-charge engines can be built to maintain a 15% coefficient of variation. However, there is certainly little reason to believe that a value much lower than 15% could be realized. On the contrary (see Appendix F, Exhaust Emissions for Chevrolet and Ford Automobiles), there is reason to believe that, in some cases at least, conventional engines exhibit engine-to-engine variability considerably exceeding this value. These facts converge in the recommendation for procurement of ten units. If variability is greater than 15% of the mean, additional vehicles could be procured as required. On the other hand, the conservative value of 15% for coefficient of variation makes it unlikely that the recommended ten vehicles would be an excessive number. If a larger coefficient of variation were assumed at the outset, however, it would dictate the procurement of a larger number of vehicles. If this estimate of the coefficient of variation proved to be too large, an unnecessary investment in vehicles would have been incurred.

It is recognized, of course, that the choice of a 95% confidence interval is somewhat arbitrary and that the specification of its half-width as 10% of the expected value is equally so. As is pointed out in Appendix I, however, these choices represent acceptable compromises between the opposing factors of increased precision and cost.

F. Hardware - Information Considerations

One aspect of an evaluation of the stratified-engine involves a consideration of the types of data that can be most expeditiously obtained with engines that have been (a) hand-built, (b) built with soft-tooling techniques, and (c) built by standard mass production techniques. A corollary problem concerns the relative resources that should be allocated to any one or all of these options.

In assessing the feasibility of mass-producing stratified-charge engines with acceptable exhaust emissions performance, the major concern is with evaluating the within-engine and between-engine variances.* Therefore, the relative usefulness of the type of engine manufacture (hand-built, soft-tooled, or production) must be judged in relation to the capability of its yielding the desired experimental data.

---

* An exceptionally high sensitivity of engine performance to changes in ambient conditions (temperature, humidity, and pressure) obviously could not be ignored.

Extracting the components of variance associated with the engines (between and within) involves accounting for all other (non-engine) sources of variance. This differentiation can be done by many repeated tests upon a single test specimen engine (presuming the within-engine variance is small or separately established). There is no basis in fact to suppose that the ability of an engine to provide consistently reproducible performance upon repeated testing should depend upon the type of manufacture involved. On the other hand, the between-engine variance would be expected to strongly depend upon manufacturing tolerances achieved which are in turn dependent upon the type of manufacture.

The hand-built engine is used to verify and prove new concepts, principles and designs. Tolerances between engines would be expected to be poorer than in either the soft-tooled or production cases. Further, hand-built engines generally tend to be in a continuous state of evolution with each succeeding unit differing in some way from the preceding one. This is the case with existing hand-built engines so that even an assessment of the between-engine variability of hand-built engines is not possible unless the family tested consists indeed of a number of carbon copies.

The major difference between soft-tooling and regular production tooling lies with the lower initial cost and poorer durability of the former. As far as tolerances are concerned, those obtained with soft-tooling should be equivalent to those achieved with normal production tooling.

It is trivial to note that the most reliable data would be obtained by making tests on samples selected from an actual population of production engines. Both economic and time considerations would dictate that a commitment to even limited tooling of the stratified-charge engine would not be made until compelling evidence of a high probability of success had been obtained. Such evidence is presently not available.

Based on the considerations that have been cited in the preceding discussion, the following recommendations are made:

(1) Soft-tooled engines will provide the more meaningful test results. Since the decision between the competing FCP and TCP engines must precede any commitment to soft-tooling, a sufficient quantity of identical hand-built engines should be procured in the interim to facilitate durability testing.

(2) When procurement is initiated for soft-tooling, it should be sufficient to fabricate a minimum of ten and possibly up to thirty 4-cylinder, stratified-charge engines with tolerances equivalent to mass production standards.

(3) Procurement should be initiated for ten 4-cylinder, stratified-charge engines built with the soft-tooling specified in item (2) above.

APPENDIX A - SUMMARY OF A LITERATURE SURVEY OF THE STRATIFIED CHARGE ENGINE CONCEPT

L. Bogdan

OBJECTIVE

The purpose of this report is to consolidate technical information, as gleaned from a survey of the literature, on the critical design factors, components and parameters that affect the operational performance and exhaust emissions of the stratified charge engine (SCE). Compromise tradeoff possibilities are discussed with their concomitant consequences. Problem areas are delineated.

INTRODUCTION

Historically, the concept of the stratified charge engine dates back over forty years. The initial concept evolved from efforts to control detonation ("knocking") in internal-combustion, spark-ignited engines by making the end gases in the cylinders incapable of autoignition. One method of attaining this objective is to separate or stratify the fuel-air content of the cylinder in such a way that the end gas lacks sufficient fuel to support auto-ignition. Engines employing this principle are referred to as stratified charge engines and, in general, achieve stratification by fuel injection late in the compression stroke.

The principles of the stratified charge concept have been developed more fully by the Texas Company with exploratory research commencing in the late 1940's and by the Ford Motor Company which first built a successfully operating laboratory engine in 1960. While significant research by other organizations has also been conducted, the two aforementioned companies dominate the field currently and have each fabricated a quantity of prototype, handbuilt, multicylinder stratified charge engines. This paper accordingly will restrict its scope to a consideration of the two engine types based on the Texaco Combustion Process (TCP) and the Ford Combustion Process (FCP).

The SCE, often referred to as a hybrid engine, combines essential elements of both the spark-ignition and compression-ignition internal combustion engines and also enjoys some of the better operational consequences of both design philosophies. Fuel injection with attendant stratification facilitates lean operation resulting in significant fuel economies and the use of high compression ratios, without constraints on fuel octane requirements, permits attainment of efficient power output relative to a carbureted, homogeneous charge engine. Positive ignition, on the other hand, results in a smooth combustion process that permits a compact, lightweight engine design combining low cost with excellent cold-starting characteristics as contrasted with the diesel engine.

These salient features of the SCE attracted the interest of the U.S. Army Tank Automotive Command (USATACOM) which is presently supporting the development of the TCP and FCP concepts with the objective of realizing an SCE replacement for the standard 4-cylinder L-141 jeep engine. The goals of the program are to develop an SCE engine having equivalent performance characteristics to the L-141 but with significantly better fuel economy and an ability to operate satisfactorily using a wide range of military fuels. A further objective is a maximum measure of parts commonality with the standard L-141 engine.

Early tests conducted on the TCP and FCP versions of the L-141 engine indicated that low exhaust emissions were also characteristic of this design. Consequently the Environmental Protection Agency (EPA) is contributing financially to the support of the USATACOM development but with the express interest of ascertaining the technological feasibility of the SCE concept meeting the 1976 emissions and durability standards in passenger vehicle applications.

DESIGN CONCEPTS OF THE SCE

The TCP and FCP concepts have been aptly called controlled combustion and programmed combustion processes. Before discussing the broader aspects of the design of the two approaches, which physically and conceptually are very similar, it is well to contrast the differences as exemplified by the stated objectives of the two companies. It must be borne in mind that Texaco is not an engine manufacturer and hence has stressed concept understanding in its research efforts. Further, Texaco's engine development effort appears to be circumscribed by the USATACOM L-141 engine development program whereas Ford, besides its role in the L-141 project, has built and tested large-size V-8 SCE's (up to 534 in.^{3} displacement; the L-141 has 141 in.^{3} displacement) with a view to passenger vehicle and (possibly) truck applications.

Texaco objectives for the TCP engines, as conditioned by USATACOM requirements are:

- good fuel economy
- broad fuel tolerance
- good durability

Ford objectives for the FCP engine include:

- fuel economy
- smoke-free power output equivalent to carbureted engine operated on regular grade gasoline
- cold start capability without special aids
- parts commonality with carbureted engine

- speed/torque characteristics approximating carbureted engine for drive train/exhaust system commonality
- low emission characteristics
- engine smoothness consistent with passenger vehicle standards

The stratified charge engine, as exemplified by the TCP and FCP, may be described a spark-ignition, internal combustion system in which fuel-air content of the cylinder is so stratified, using fuel injection, that the richest mixture is concentrated in the region of the spark plug electrodes. Thus ignition of the charge is readily initiated and operation at lean overall fuel/air ratios at part load conditions is feasible. Inlet air throttling is not required with the power level regulated by the duration of the fuel injection.

Studies have established the fact that one of the key aspects of the stratification process, which is essential to the successful realization of a useful engine, is the development and maintenance of a suitable swirling motion of the air within the cylinder. In fact, implementation of the SCE concept is entirely dependent upon proper coordination of three primary parameters: air swirl, fuel injection and positive spark ignition. Each of these parameters will be examined in turn in terms of physical implications of design and relation to performance. It will be appreciated that exploratory experimentation underlying the development of useful design data has been empirical in nature using trial-and-error type procedures.

AIR SWIRL

Air swirl rate (rotational velocity of the air relative to crankshaft angular velocity) has a direct bearing on the combustion duration and efficiency serving as it does to direct the injected fuel to the flame front and to control the evaporation and dispersion of the fuel. The realization of a suitable swirl rate with an attendant high volumetric efficiency over the useful operating range of the engine constitutes one of the primary design variables.

A swirling airflow geometry free of counterflows and irregularities is conducive to consistent engine operation from cycle to cycle. To achieve this goal much experimental effort has been devoted to define geometries for: directional intake ports, stationary shrouds around intake valves and cam profiles and valve lifts. All of these factors serve to induce the initial, low-rate swirl that exists during the early phases of the compression stroke. The necessary high swirl rates are achieved by means of a recessed, centrally located, combustion cup or chamber which is an integral part of the piston crown. As the piston approaches the top of its compression stroke, the swirling air is constrained to within the confines of the smaller-than-bore diameter of the cup and the swirl rate increases in consequence of the conservation of angular momentum. Radial dimensions of the cup affect swirl rate while the volume determines the compression ratio.

The cup dimensions and geometry have a significant impact on the combustion process for reasons that are not completely understood. Test data have shown that the ratio of the area of the cup aperture to the bore area represents a sensitive parameter. If the ratio is large, fuel economy is poor; when the ratio is small, combustion roughness occurs. A typical ratio is 0.42 which suggests that charge compactness is necessary during combustion to provide a high rate of heat release.

A wide range of cup geometries has been explored: truncated cone, cone cup, double-cone cup, etc. The main effect reported is on fuel economy and combustion harshness with the double-cone cup most successful. This performance is attributed to the sharp-edged surface discontinuities at the cone-cone and cone-cup intersections which create flow turbulence and mixing resulting in accelerated heat release with smooth combustion. This assumption has been supported by tests which demonstrated poorer performance when the cup edges were rounded.

To a degree, the cup-in-piston concept is a consequence of the necessity to achieve adequate swirl rate within the restrictions imposed by the oversquare bore-to-stroke ratio of the basic L-141 engine which represents an unfavorable geometry in that respect. In addition, in SCE applications, large bore-to-stroke ratios have been determined unfavorable from the standpoint of good thermal efficiency and low hydrocarbon (HC) emissions.

To fully explore and exploit the potential of a stratified charge engine concept, it seems unreasonable to fetter the design of prototype models by slavish insistence upon maximum parts commonality with an existing engine design optimized for different combustion processes. By permitting the bore-to-stroke ratio to be a design variable, it appears that improvements in air swirl formation, engine thermal efficiency and HC emissions could be realized.

To the extent that such artifacts as the combustion cup with its surface discontinuities (ridges) are crucial to successful operation of the SCE with the L-141 configuration, it is possible that combustion deposits, by altering critical shapes, could unfavorably affect durability. Thus an improved bore-stroke design could also result in lesser deterioration of engine performance (including emissions) with accumulated mileage.

FUEL INJECTION

Optimization and the consistent reproducibility of the fuel injection process represents the very crux of the stratified charge engine operation. In this regard the fuel injector is the critical mechanical component since the mixture formation is controlled primarily by the characteristics of the fuel spray.

A large amount of empirical study has been expended in determining the proper location for the fuel injector, the angle of the fuel injection (relative to the bore centerline), fuel injection pressure, fuel injection rate and fuel spray characteristics (spray angle). The objective has been to achieve satisfactory operation over the entire speed-load range capitalizing upon the potential of lean operation of the stratified charge principle at partial loads while being able to realize the full output capability of the homogeneous charge principles at maximum loads.

Tests have shown that at light load operation it is desirable to confine the fuel in the proximity of the center of the bore while at high loads good air utilization requires that more fuel be directed radially outward. Because of mutual interaction, a compromise is effected between injector orientation and injector spray angle. Basically, a wide angle fuel spray is required to achieve low penetration. This condition may lead to cylinder head or wall wetting by the fuel resulting in rough combustion. A narrow angle spray increases penetration and hence dispersion with misfire or poor combustion characteristics at light loads.

Injection pressure is another factor affecting fuel penetration but more significantly is its role in the atomization of fuel. Too low injection pressures appear to produce poor atomization with resultant losses in fuel economy and increases in HC emissions. Injector opening pressure typically range from 300 to 450 psi since higher pressures do not seem to produce any advantages.

The fuel injection rate is constrained to a very low level so that the duration of the injection is extended to facilitate operation and render timing coordination less critical. Other benefits accruing from low fuel mass injection rates are: lesser disturbance of air swirl pattern, better fuel evaporation and improved fuel-air mixing.

Much effort has been expended by Ford on the design, development and evaluation of injector nozzles to achieve reliability, consistency of performance and freedom from fouling. Designs inducing valve vibration and rotation have been favored since they tend to: promote atomization of the full, minimize carbon buildup, provide better valve seating and maintain better spray concentricity. These designs seem to be susceptible to cyclic variations in fuel delivery resulting from mechanical resonances in the valve. Ford has similarly developed specialized, multicylinder injection pumps to optimize the entire fuel injection system.

Texaco has apparently been able to successfully adapt standard diesel injection pumps and nozzles to their TCP engines. Ford's experience led them to reject commercially available components as being unsuitable.

POSITIVE SPARK IGNITION

Conceptually, the FCP and TCP systems employ the conventional coil, distributor, spark plug system. In practice, the differences from convention are significant. Rather detailed information on the FCP developmental efforts in this regard is given in the open literature.

Ford conducted extensive studies to determine the optimum spark plug location for consistently reliable fuel mixture ignition over the operating range of the engine. In the early designs, this requirement was best realized with the gap in the center of the fuel spray near the bore centerline and about one-half inch below the cylinder head face. The basis for this location seemed to be that the central core of the mixture cloud rotates as a solid body without appreciable dispersion and hence is an ideal point for flame initiation.

One problem with this gap location was that long electrodes on the spark plug were required. In addition, due to the asymmetrical electrode geometry, the orientation of the plug was important and a fixed-orientation plug installation was mandatory. A potential difficulty with long electrodes is their tendency to overheating with resultant preignition.

In the latest (PROCO) version of the engine which makes use of air throttling and exhaust gas recirculation, the center of the spray mixture was deemed to be richer than desirable for flame initiation, especially under medium to heavy load conditions. Consequently a new, short-electrode spark plug has been developed with the gap located at the edge of the fuel spray (above the spray centerline). This location has had several beneficial effects such as: increased tolerance to EGR, lower specific fuel consumption, and decreased levels of HC and CO emissions. A slight increase in NO emission was observed but this condition can be corrected by changes in EGR rate and/or ignition timing.

In the FCP engine, spark gap erosion problems have been severe (a gap growth rate four to five times that experienced in carbureted engines). This erosion is a consequence of the high-energy requirements of the ignition system to preclude part-power misfires and the high electrode temperatures. After studying various alternatives, Ford finally resorted to riveted precious metal (gold-palladium) inserts in the gap area.

Texaco makes no reference to the spark plugs used in its TCP engines but rather emphasizes development of a special transistorized ignition system with the following spark characteristics: high energy output, controlled duration at essentially constant voltage, restrike capability and very fast rise time. This system is claimed to permit operation in a satisfactory manner at leaner mixtures than is possible with a conventional ignition system.

INJECTION - IGNITION TIMING

The coordination of the relative timing among fuel injection, spark ignition and crank angle on the compression stroke is vital to the achievement of proper combustion control in the stratified charge engine. In the FCP and TCP versions, different philosophies have been adopted with respect to timing resulting in different operational benefits. The differences are made manifest in the following discussion.

The Ford system utilizes independent advance-retard timing schedules for spark and injection as functions of engine speed and load. At light loads a well-stratified mixture is essential to achieve satisfactory ignition. To achieve this end, the fuel injection occurs late in the compression stroke. A small, stratified mixture cloud is formed in the proximity of the centrally-located spark gap. The mixture cloud gradually diffuses in the swirling airflow pattern while remaining centered on the spark gap. Ignition is initiated during the period of injection. With increased loads, injection timing is advanced to allow more time for fuel dispersion/evaporation for faster and more complete combustion. At the same time, spark timing is retarded as combustion rate increases with richer mixtures. With increasing engine speeds, both injection and ignition timing are advanced so that the mixture cloud formation and the combustion process can accommodate the higher piston speeds.

Ford's timing schedule provides excellent air utilization at high loads and thus permits the realization of full power output without the smoke problems normally associated with heterogeneous charge engines under these operating conditions. On the other hand, the engine becomes susceptible to preignition and detonation difficulties with attendant limitations on the useful ranges of compression ratios and fuel octane ratings.

The relative phasing of the injection and ignition timing is critical and the magnitudes are large. For example, injection timing covers a range of approximately 40° with speed and 60° with load. Similarly, the respective values for spark timing are 15° and 25°. To achieve and maintain this timing schedule over the useful life of the engine by means of independent adjustments would require exceptionally knowledgeable and skilled maintenance personnel. Ford is therefore opting for a unitized injection/ignition system incorporating the proper timing schedule.

Texaco literature lacks precise detail concerning the injection/ignition timing schedule used in the TCP. From basic considerations, it is conjectured that the fuel injection timing relative to crank angle must approximate that of the FCP. Ignition, however, appears to be timed so that the first increments of the injected fuel are <u>always</u> ignited. This ignition time precludes the possibility of preignition and eliminates the possibility of detonation regardless of the octane rating of the fuel. The realization of these possibilities has been amply demonstrated by actual in-vehicle tests. The consequence of these real advantages is that the air utilization at high load conditions becomes poor so that the maximum output is exhaust smoke limited.

The relative freedom in selecting compression rates makes the TCP a ready candidate for turbocharging. Dynamometer tests of a turbo-charged TCP engine (without changes to the injection system or valve train) have shown no misfire or knock, no adverse effects on part-power fuel economy but with a good improvement in high load performance. Vehicle tests demonstrated improved fuel economy (relative to a naturally aspirated engine) at all but the very low operating speeds (for all fuels tested).

The TCP and FCP demonstrate the extent to which injection/ignition timing may be used to attain different operational objectives; in brief, a broad fuel tolerance for the TCP and excellent, smoke-free full load operation for the FCP. Considerable latitude appears to exist here to explore intermediate timing schedules to achieve a "best of both worlds" performance from a stratified charge engine.

PERFORMANCE

A. Four Cylinder SCE

Both the TCP and the FCP have fulfilled their different design objectives while equalling or exceeding the torque/speed characteristics of the standard carbureted L-141 engine. Good cold weather starting has been achieved and the ability of immediate after-start driveaway without the need for engine warm-up is facilitated by fuel injection. Excellent throttle response, bordering on harshness, has been demonstrated. Over-the-road tests provided excellent proof of fuel economy, relative to the standard engine, of 40% to 70%.

The unthrottled SCE has poor idling characteristics which may be resolved either by employing high engine idle rpm or by incorporating air throttling that is only operable at idle. While the early prototypes were unthrottled, the trend in the latest versions is to incorporate some throttling over the entire spectrum of engine operating conditions.

In common with standard carbureted engines, the addition of emission control devices to the SCE has resulted in a loss of performance (power) and fuel economy. This is especially true in the use of exhaust gas recirculation (EGR) to control oxides of nitrogen in the exhaust.

B. Eight Cylinder SCE

Ford has converted, on its own funding, 430 CID and 534 CID V-8 engines to the FCP configuration using the nominal valve train and camshafts found in these engines so that maximum efficiency was not realized. In passenger vehicle application, the smaller engine produced gains in fuel economy of about 30% over a wide range of operating speeds (30-70 mph). Acceleration and passing performance was found to be somewhat deficient despite dynamometer test data which indicated power equal to (or better than) the basic standard engine. Serious idle problems were encountered which necessitated idle throttling. The high peak instantaneous torques of the FCP resulted in noise, vibration and harshness characteristics considered unacceptable in the luxury-type cars normally using this size engine. Improved engine mounting would be required to alleviate these latter difficulties.

The larger engine (a truck engine) was converted principally to obtain scaling data (Ford does not consider FCP engines as candidates for truck applications). Compared with the 430 CID engine, test data on the 534 CID unit showed that: compression ratio had to be reduced to eliminate spark knock, misfire was more prevalent, severity of spark gap erosion increased, high-speed fuel economy was poorer and control of hydrocarbon emissions was more difficult.

In terms of scaling effects, it was concluded that: (a) increased bore diameter necessitates use of lower compression ratios which increase tendency to misfire and require higher energy ignition and (b) increased bore diameter increases the volume of very lean mixture which isn't burned and hence results in increased HC emissions. Ford concluded that fuel consumption and exhaust emission data show that good correlation exists between single and multicylinder engines.

EMISSIONS

In homogeneous charge, internal combustion engines, a large measure of the exhaust emission is associated with wall-wetting of the intake manifold by fuel, flame quench at the walls of the combustion chamber and a deficiency of air. Emissions from the SCE are smaller to the extent these conditions are eliminated.

In the SCE the principal source of hydrocarbon emissions is the stray droplets of fuel that escape from the fuel spray and the overly-lean mixture in the peripheral regions that does not burn or burns incompletely. Unburned hydrocarbon content in the exhaust decreases when injection timing is retarded and/or spark timing is advanced since the time for fuel dispersal prior to flame front arrival is decreased. Intake throttling also has a favorable effect in decreasing HC emission since the fringe area leanness is reduced thus facilitating outward penetration of the flame producing more thorough combustion.

With an excess of air available at all operating conditions, carbon monoxide emissions have assumed negligible proportions. If additional throttling is to be added to the basic SCE to realize other engine operational benefits, it is possible that CO emissions will be unfavorably affected.

The oxides of nitrogen in the exhaust of an SCE are approximately equivalent to those found in the conventional counterpart engine. As in conventional engines, EGR has been found to be an effective control to reducing NO_x.

The stratified charge engine is at a definite disadvantage (relative to homogeneous charge engines) in the matter of aldehyde emissions. Aldehydes are believed to be the cause of the distinct odor associated with

SCE exhaust gases. This odor is unpleasant to the sense of smell and in sufficient concentrations irritates the eyes and nose. Partial throttling has been found to reduce the odor which also is affected by ignition and spark timing. With a compromise throttling and timing situation, a catalytic muffler might be required for overall control. A potential problem is the relatively low level exhaust gas temperature found in the SCE. As a consequence the catalyst efficiency could be impaired by low operating temperatures.

Recent emissions data taken by EPA on the 4-cylinder FCP installed in a military jeep and equipped with a properly-tuned EGR system and an exhaust catalyst have shown its ability to meet Federal 1976 emissions standards. To achieve these levels, however, EGR was used at wide open throttle without simultaneous fuel enrichment. This mode of operation results in about a 30% loss in power with resultant poor acceleration and misfire at high speeds.

Mutually conflicting requirements are therefore seen to exist between achieving low HC/CO emissions and low NO_x emissions. A similar situation obtains between low NO_x emissions and engine full load performance Obviously these conditions represent an area where trade-off and compromise between emissions, engine performance and engine size (to make up for loss of performance due to the use of emission control procedures and devices) is both possible and desirable.

PROBLEM AREAS

As viewed from the standpoint of whether the SCE has the ability to fulfill the role of a satisfactory passenger vehicle power plant and still meet the Federal emission requirements for 1976, two principal questions will need to be resolved: (a) the feasibility of mass-producing the SCE to the necessary tolerances and (b) the capability of maintaining low emissions performance over the necessary 5-year/50,000 mile endurance period.

Production feasibility is concerned primarily with two crucial components; the fuel injection pump and the fuel injection nozzles (especially in the case of the FCP). The feasibility of mass producing a satisfactory unitized injection/ignition system is also a vital consideration.

Durability considerations (unique to the SCE - deterioration of emission control systems is a problem common to all internal combustion engines) include injector coking, fouling or inability to render consistent performance, failure of the injection pump to render consistent performance and spark plug gap erosion.

Exhaust gas odor, characteristic of most heterogeneous charge engines, will need to be controlled if the SCE is to be a thoroughly acceptable vehicle power plant. In the multifuel SCE, smoke emissions at high load conditions represent a problem. While different methods of attack have been put forward as possible solutions, no successful demonstrations of satisfactory control are known.

RECOMMENDATIONS

Within the limitations of the literature search conducted, it appears that all of the multicylinder SCE engines that have been built and tested to date represent adaptations, to a lesser or greater degree, of an existing carbureted-engine design. The end objective has generally a maximization of parts commonality. As discussed in the preceding sections, this approach imposes restraints in design that preclude the full potential of the stratified charge combustion process to be achieved. It is therefore

recommended that in an early stage of the evaluation process of the pr
available family of SCE models, if the producibility and durability que
are satisfactorily answered, that an engine fully optimized to the strat
charge process and configured for low emissions be designed, built an
evaluated.

BIBLIOGRAPHY

1. Taylor, C. F., "The Internal Combustion Engine in Theory and Practice - Volume II: Combustion, Fuels, Materials, Design," The M.I.T. Press, Cambridge, Massachusetts, 1968.

2. Bishop, I. N. and Simko, Aladar, "A New Concept of Stratified Charge Combustion The Ford Combustion Process (FCP)," SAE Paper 680041, January 1968.

3. Mitchell, E., Cobb, J. M. and Frost, R. A., "Design and Evaluation of a Stratified Charge Multifuel Military Engine," SAE Paper 680042, January 1968.

4. Summerson, W. A. and Mitchell, E., "Small Military Engine Stratified Charge Combustion Yields Multifuel Capability," Automotive Industries, December 1967, pp. 71-76.

5. Haman, A. C., Cheklich, G. E. and Kaupert, A. W., "Review of the USATACOM Hybrid Combustion Engine Program," U. S. Army Tank Automotive Command, no number, August 1970.

6. "The Texaco Controlled-Combustion System A Stratified Charge Engine Concept," Texaco Research Center, no number, October 1970.

7. "The Texaco Transistorized Ignition System," Texaco Research Center, no number, November 1970.

8. "Engineering Know-How in Engine Design Part 19," Society of Automotive Engineers, Inc., SP-365, June 1971.

9. "Control Techniques for Carbon Monoxide, Nitrogen Oxide, and Hydrocarbon Emissions from Mobile Sources," U.S. Department of Health, Education and Welfare, AP-66, March 1970.

10. Simko, A., Choma, M.A., and Repko, L.L., "Exhaust Emission Control by the Ford Programmed Combustion Process - PROCO," SAE Paper 720052, January 1972.

APPENDIX B ANALYSIS OF EMISSIONS TEST DATA ON LOW-EMISSION FCP ENGINE INSTALLED IN M151 VEHICLE

D. J. Schuring

INTRODUCTION

The Ford Motor Company, under contract with U. S. Army Tank Automotive Command, adapted its stratified-charge combustion process (FCP) to the L-141 engine of the M 151 vehicle (jeep). Ford developed two types of engines -- a low-emission engine and a best-economy engine -- and installed them in four vehicles, three jeeps and one two-by-four post office truck. Of the four vehicles, only vehicle No. 2 (2M5733), a jeep with a low-emission FCP engine, is of interest here. The engine of this vehicle was equipped with exhaust gas recirculation (EGR with heat exchanger) and a PTX-6 catalyst.

The vehicle has been repeatedly tested on a chassis dynamometer at Ford and EPA using the 1975 testing and sampling procedure (CVS-CH with three bags). (This procedure could not be completely followed, though; the engine lagged significantly behind the prescribed acceleration rates during the 192 215 sec. portion and the 452 470 sec. portion of the driving cycle because, in order to keep emissions low at high acceleration rates, the EGR was not closed and fuel enrichment was not applied at wide open throttle.) During the tests, a few changes of the engine and the emission-control hardware were made, such as replacement of a faulty catalyst, adjustment of the EGR system, and adjustment of the deceleration fuel cutoff linkage.

The data of all tests were analysed with the objective of clarifying the following questions.

What are the average values and the variance of the three pollutants of the given engine?

Did the changes made during the tests significantly influence the emission level?

To what extent did the tested vehicle meet the 1976 emission standards?

Data

The results of emission tests performed between July 22 and September 29 (1971) were documented in three Ford Engineering Progress Reports[1-3] and in computer printouts of EPA[4]. Tables 1 through 6 list all pertinent data.

Data set No. 1 (Table 1)	4 tests run between July 22 and July 29 at Ford. Faulty PTX-6 palladium catalyst may have accounted for high HC and CO data and for large data spread. The Ford Progress Report remarks that the very high CO-value of 5.2 gr/mi (July 26) was due to low barometric pressure at the day of testing. Since a significant correlation between barometric pressure and emission level could not be established, however, (as reported elsewhere) we tend to ascribe the unusually high CO value to other unknown reasons. In any event, the CO value of 5.2 gr/mi was disregarded in further data evaluation.
Data set No. 2 (Table 2)	5 tests between August 13 and 20 at Ford with a new PTX-6 platinum catalyst.
Data set No. 3 (Table 3)	5 tests between August 30 and September 3 at EPA, Ypsilanti.
Data set No. 4 (Table 4)	4 tests between September 7 and 9 at EPA, Ypsilanti. Ford personnel adjusted accelerator linkage to prevent closing of EGR at wide open throttle.
Data set No. 5 (Table 5)	5 tests between September 9 and 14 at EPA, Ypsilanti. Ford personnel adjusted the deceleration fuel cutoff setting.

Date 1971	HC-FID g/mi	CO-IR g/mi	NO_x-Cl g/mi	EGR Orifice diam.	Vacuum EGR Setting	Inertia Weight	Inlet Pres. in Hg	Barom. Pres. in Hg	Temp. °F	Humidity %
7-22	.38	1.6	.34	0.438				29.32		
7-23	.32	1.9	.35					.22		
7-26	.44	5.2	.28					.05		
7-29	.51	1.8	.28					.28		
mean $\bar{x}$	.413	2.63 (1.77)*	.313							
Std Dev s	.0814	1.73 (.153)*	.0377							
$s/\bar{x}$ %	≈ 20	≈ 66 (≈9)*	≈ 12							

TABLE 1 Emission test data of Vehicle No. 2 (2M5733) tested at Ford, Dearborn
Source of Information: Ford Engrg. Progress Rpt. 14 (July 1971)
Vehicle mileage 26,949, FCP engine mileage 3502. Apparently tests were run with faulty PTX-6 palladium catalyst

* without CO = 5.2 g/mi.

Date 1971	HC-FID g/mi	CO-IR g/mi	NO_x-Cl g/mi	EGR Orifice diam.	Vacuum EGR Setting	Inertia Weight	Inlet Pres. in Hg	Barom. Pres. in Hg	Temp °F	Humidity %
8-13	.28	.54	.42	.375				29.32		
8-16	.33	.32	.37			3000		.46		
8-17	.35	.21	.35		.50			.40		
8-18	.27	.14	.32	.404		2000		.28		
8-20	.27	.42	.41		.25	3000		.16		
mean $\bar{x}$	.300	.326	.374							
Std. Dev. s	.0374	.160	.0416							
$s/\bar{x}$ %	≈12	≈49	≈11							

TABLE 2 Emission test data of Vehicle No. 2 (2M5733) tested at Ford, Dearborn
Source of information: Ford Engrg. Progress Rpt. 15 (August 1971)
Vehicle Mileage 27,055, FCP engine mileage 3,566. Tests were run with new PTX-6 platinum catalyst.

Date 1971	HC-FID g/mi	CO-IR g/mi	NO_x-Cl g/mi	EGR Orifice diam.	Vacuum EGR Setting	Inertia Weight	Inlet Pres. in Hg	Barom. Pres. in Hg	Temp °F	Humidity %
8-30	.41	.98	.45				26.20	29.36	81	46
8-31	.37	.72	.39				.30	.46	72	56
9-1	.46	.84	.39			2750	.27	.42	76	60
9-2	.46	.82	.37				.21	.38	76	73
9-3	.33	.81	.49				.22	.37	75	75
mean $\bar{x}$	.406	.834	.418							
Std. Dev s	.0568	.0937	.0624							
$s/\bar{x}$ %	$\approx$ 14	$\approx$ 11	$\approx$ 12							

TABLE 3 Emission test data of Vehicle No. 2 (2M5733) tested at EPA, Ypsilanti
Source of information: EPA Computer Data
Vehicle mileage 27,155, FCP engine mileage 3670 (estimated)

Date 1971	HC-FID g/mi	CO-IR g/mi	NO_x-Cl g/mi	EGR Orifice diam	Vacuum EGR Setting	Inertia Weight	Inlet Pres. in Hg	Barom. Pres. in Hg.	Temp °F	Humidity %
9-7	.37	.92	.21			2750	26.13	29.32	74	74
9-7	.47	1.17	.34			2750	27.66	.25	81	64
9-8	.39	.92	.32			3000	26.12	.25	83	61
9-9	.32	.83	.31			3000	.19	.30	76	65
mean $\bar{x}$	.388	.960	.295							
Std. Dev s	.0624	.146	.0580							
$s/\bar{x}$ %	≈ 16	≈ 15	≈ 20							

TABLE 4 Emission test data of Vehicle No. 2 (2M5733) tested at EPA, Ypsilanti
Source of information: EPA Computer data
Vehicle mileage 27,227, FCP engine mileage 3730 (estimated)
EGR linkage adjusted by Ford personnel to prevent EGR cutoff at WOT

Date 1971	HC-FID g/mi	CO-IR g/mi	NO_x-Cl g/mi	EGR Orifice diam.	Vacuum EGR Setting	Inertia Weight	Inlet. Pres. in Hg	Barom. Pres. in Hg	Temp. °F	Humidity %
9-9	.27	.58	.28				26.16	29.38	81	50
9-10	.30	.55	.29			3000	.07	.18	74	64
9-13	.32	.78	.26				25.94	.03	70	63
9-13	.31	1.20	.25				.91	.00	72	58
9-14	.40	1.24	.35				.91	28.99	70	62
mean $\bar{x}$	.320	.870	.286							
Std. Dev s	.0485	.332	.0391							
$s/\bar{x}$ %	≈ 15	≈ 38	≈ 14							

TABLE 5 Emission test data of Vehicle No. 2 (2M5733) tested at EPA, Ypsilanti
Source of information: EPA Computer data
Vehicle mileage 27,279, FCP engine mileage 3790 (estimated)
Accelerator linkage adjusted by Ford personnel to prevent closing of EGR at WOT
Deceleration fuel cutoff adjusted by Ford personnel

Date 1971	HC-FID g/mi	CO-IR g/mi	NO_x-Cl g/mi	EGR Orifice diam.	Vacuum EGR Setting	Inertia Weight	Inlet. Pres. in Hg	Barom. Pres. in Hg	Temp. ^{o}F	Humidity %
9-27	.34	.36	.31	.404	.50			29.29		
9-29	.25	.33	.30	.404	.50			29.36		
mean $\bar{x}$	.295	.345	.305							
Std. Dev. s	.0636	.0212	.00706							
$s/\bar{x}$ %	$\approx$ 22	$\approx$ 6	$\approx$ 2%							

TABLE 6 Emission test data of Vehicle No. 2 (2M5733) tested at Ford, Dearborn
Source of Information: Ford Engrg. Progress Rpt. 16 (September 1971)
Vehicle mileage 27,406, FCP engine mileage 3917

Data set No. 6 (Table 6) 2 tests between September 27 and 29 at Ford.

Methodology

Each of the six data sets was treated as a homogeneous sample (although small engine changes did occur during some test series as evidenced by alterations recorded in Table 2). The hypothesis was then tested whether all data sets could be considered samples of one and the same grand population.

First, the means, $\bar{x}$, and the standard deviations, s , were computed for each of the three exhaust emission gases and for each of the six sets of data, Table 1 through 6. (The coefficient of variation, $s/\bar{x}$, was added in order to indicate the data spread in relation to the mean.) Since $\bar{x}$ and s are sample means and standard deviations, we first tested whether all sample standard deviations of a particular exhaust emission gas could be considered estimates of the same population standard deviation. Variance ratios of all possible pairs of the 6 x 3 data sets were computed and compared with pertinent $F_{N_i-1,\ N_j-1}$ values on a 95% and 99% significance level.

N_i number of tests of data set i

N_j number of tests of data set j

$i \neq j$

The results are listed in Table 7, columns "standard deviation" Except for one weakly significant standard deviation (CO data set No. 3 versus No. 5), all standard deviations could be considered estimates of the same true population standard deviation, which was then estimated by computing the average values of all standard variations* for the three gases (CO data set No. 5 omitted).

average standard deviation of HC ; s_{hc} = 0.0578 gr/mi

average standard deviation of CO ; s_{co} = 0.1347 gr/mi

average standard deviation of NO_x; s_{nox} = 0.0476 gr/mi

We then tested whether all means of a particular gas could be considered representing the same true population mean. For all pairs of the 6 x 3 data sets, the value

$$t = \frac{|\bar{x}_i - \bar{x}_j|}{s} \cdot \sqrt{\frac{N_i N_j}{N_i + N_j}}$$

was computed and compared with the $t_{N_i + N_j - 1}$ – value on a 95% and 99% confidence level

$\bar{x}_i$	mean of data set i
$\bar{x}_j$	mean of data set j
s	average standard deviation

Results

Table 7, columns "mean" lists the results. They show that with respect to HC, all data sets can be considered samples of the same population except for data set No. 2 and, perhaps, No. 3. Data set No. 2 is weakly (95% confidence but not 99%) different from data sets No. 1 and 3, very likely a consequence of the new catalyst installed after completion of data set No. 1. The difference between data sets Nos. 3 and 5 is very

* These standard deviations encompass both the variations of the engine and those introduced by the test apparatus and the test procedure. A quantitative separation of the two would require additional experimentation.

Test set i versus Test set j	Mean (t-test)			Standard Deviation (F-test)		
	HC	CO	NO_x	HC	CO	NO_x
1-2	*	***	-	-	-	-
1-3	-	***	**	-	-	-
1-4	-	***	-	-	-	-
1-5	-	***	-	-	-	-
1-6	-	**	-	-	-	-
2-3	*	***	-	-	-	-
2-4	-	***	*	-	-	-
2-5	-	***	* (**)	-	-	-
2-6	-	-	-	-	-	-
3-4	-	-	**	-	-	-
3-5	($\bar{*}$)	-	**	-	*	-
3-6	-	**	*	-	-	-
4-5	-	-	-	-	-	-
4-6	-	**	-	-	-	-
5-6	-	**	-	-	-	-

TABLE 7

Significance test of six sets of data

Numbers 1 to 6 in first column refer to data sets in Tables 1 to 6, respectively.

* Hypothesis of equality rejected with 95% confidence

** Hypothesis of equality rejected with 99% confidence

($\bar{*}$), ($\overset{*}{**}$) borderline cases

*** Hypothesis of equality rejected with 99.9% confidence weak (barely 95% confidence) and can possibly be neglected. Consequently, Ford data set No. 1, the EPA data set Nos. 3, 4, and 5, and the subsequent Ford data set No. 6 can be pooled into one data set (1+3+4+5+6). Another possible combination is(2+4+5+6).

With respect to CO, Table 7 reveals that data set No. 1 is distinctly different from the rest (99.9% confidence), most likely a consequence of the faulty catalyst used by Ford in set No. 1. Also, data set No. 2 is distinctly different (99.9% confidence) from the EPA data set Nos. 3, 4, and 5, apparently a consequence of the new catalyst installed after completion of tests of data set No. 1. The EPA data set Nos. 3, 4, and 5 show no differences among themselves and can be considered homogeneous. The last Ford data set No. 6, however, deviates distinctly (99% confidence) from the EPA data and conforms with the previous Ford data set No. 2. In summary, Ford data set No. 1 has to be considered as an entity by itself; the EPA data sets can be pooled (3+4+5), and the Ford data set Nos. 2 and 4 can be pooled also. With respect to NOx, we notice a distinct difference (99% confidence) between data set Nos. 1 and 3, perhaps due to maladjusted EGR linkage. The adjustment incurred a significant change (99% confidence) in NOx content as evidenced by the significant differences between EPA data set Nos. 3-4 and Nos. 3-5. The data set Nos. 4, 5, and 6 are homogeneous, proving that the EGR linkage adjustment remained effective throughout the rest of the tests.

In order to display the results of this statistical analysis in a more effective way than is possible in a tabulation, we plotted for each of the three gases (HC, CO, and NOx) and for each of the six data sets the means and their confidence limits, $\pm p$, computed using the relation $p = s \times t_n / \sqrt{n}$, where n 5 (≈ average number of degrees of freedom for two combined tests) and t = 2.5 (95% confidence for n = 5). The data are shown in Figures 1, 2, and 3. The figures illustrate clearly the impact of the three known engine changes - the replacement of a faulty catalyst by a new one after completion of test set No. 1, the adjustment of the accelerator linkage to prevent closing of EGR at WOT after test set No. 3, and the adjustment of the deceleration fuel cutoff after test set No. 4. The significance of these changes is evidenced by the difference of the means and their associated confidence limits. As long as the difference of two means is small so that their confidence ranges overlap, it cannot be considered significant. This situation, for instance, is true for the difference caused by the adjustment of deceleration fuel cutoff in Figure 1; the HC output of data set No. 6 is not significantly different from that of data set No. 5. On the other hand, the replacement of the faulty catalyst after data set No. 1 caused a significant reduction of both HC and CO, Figures 1 and 2, whereas the NOx output remained unaffected, Figure 3. We also notice the significant reduction of NOx after adjustment of the EGR system, Figure 3 data set Nos. 3 and 4. Figures 1 and 2 reveal also that the drastic reduction of HC and especially of CO emissions after installment of the new catalyst could not be maintained on the newly achieved low level. Data set No. 3 shows that HC emissions shifted back to higher levels, although they could be again reduced in subsequent tests.

To carry the investigation somewhat further, we pooled all data sets that showed no significant differences on a 95% confidence level, as indicated in Figures 1, 2, and 3, and computed their means. One could consider these means as characteristic emission levels of test vehicle No. 2.

average HC level of data sets (1 + 3 + 4 + 5 + 6)		0.371 gr/mi (with faulty catalyst)
average HC level of data sets (2 + 4 + 5 + 6)	=	0.328 gr/mi (without faulty catalyst)
average CO level of data sets (3 + 4 + 5)		0.824 gr/mi
average HOx level of data sets (1 + 4 + 5 + 6)		0.298 gr/mi

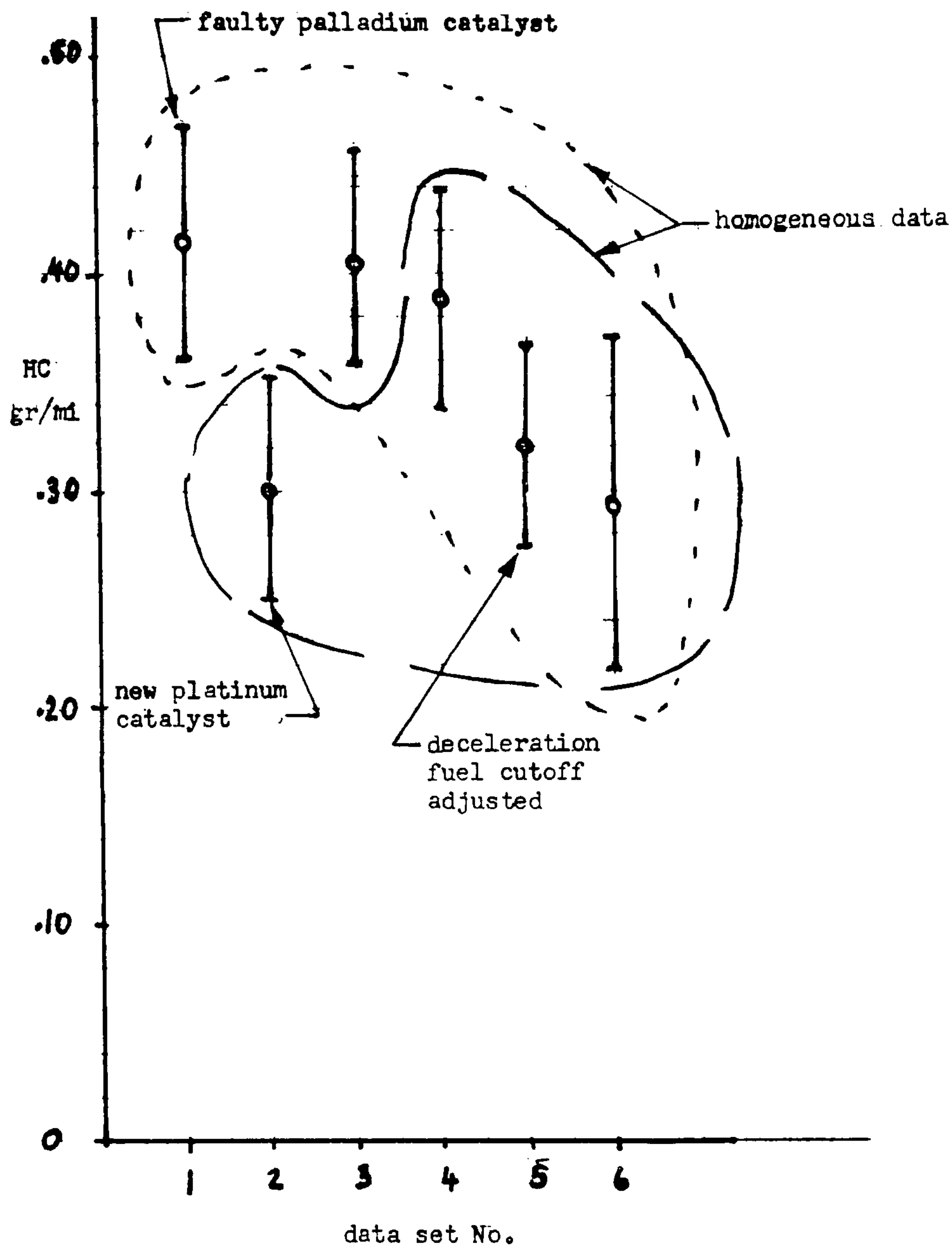


Figure 1

HC means and confidence ranges for six data sets

Data sets No. 1, 2, 6   Ford data

Data sets No. 3, 4, 5   EPA data

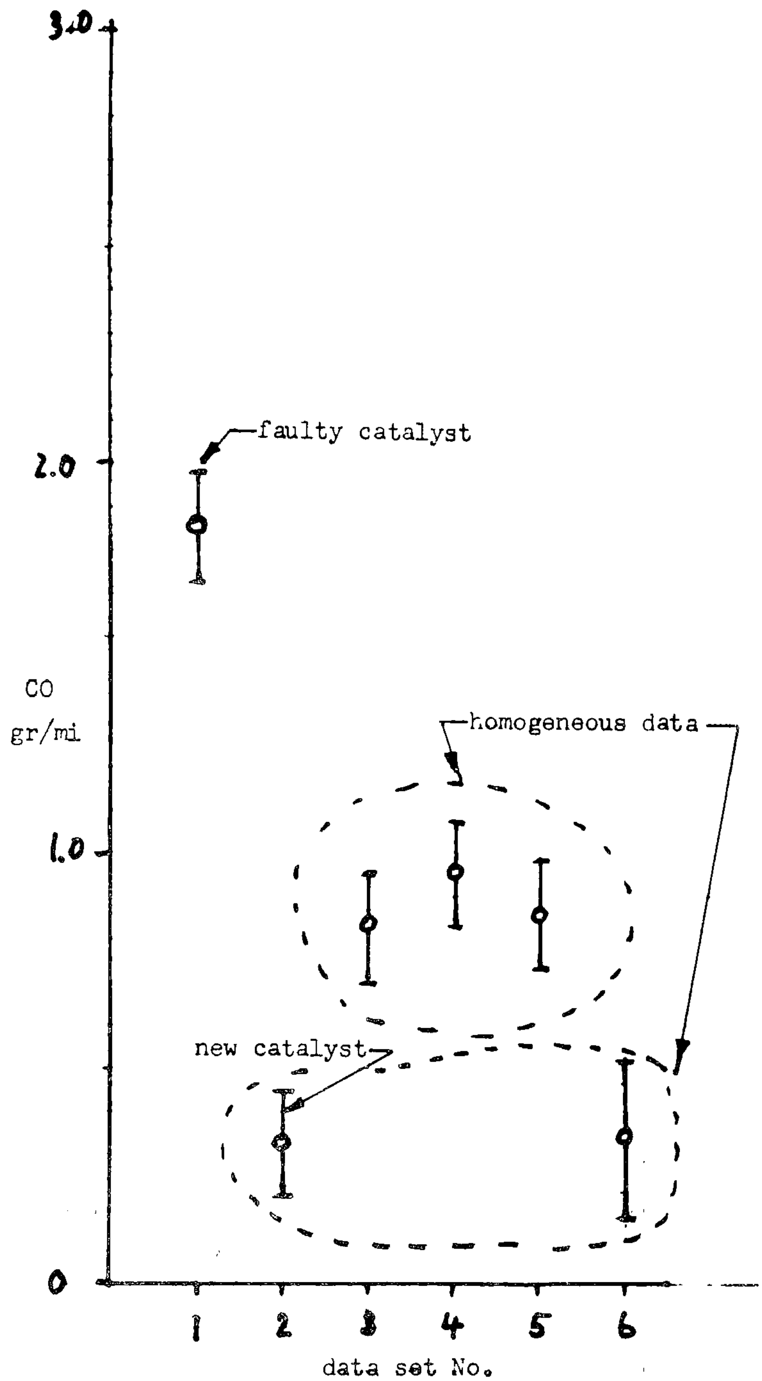


Figure 2

CO means and confidence ranges for six data sets

Data sets No. 1, 2, 6 Ford data

Data sets No. 3, 4, 5 EPA data

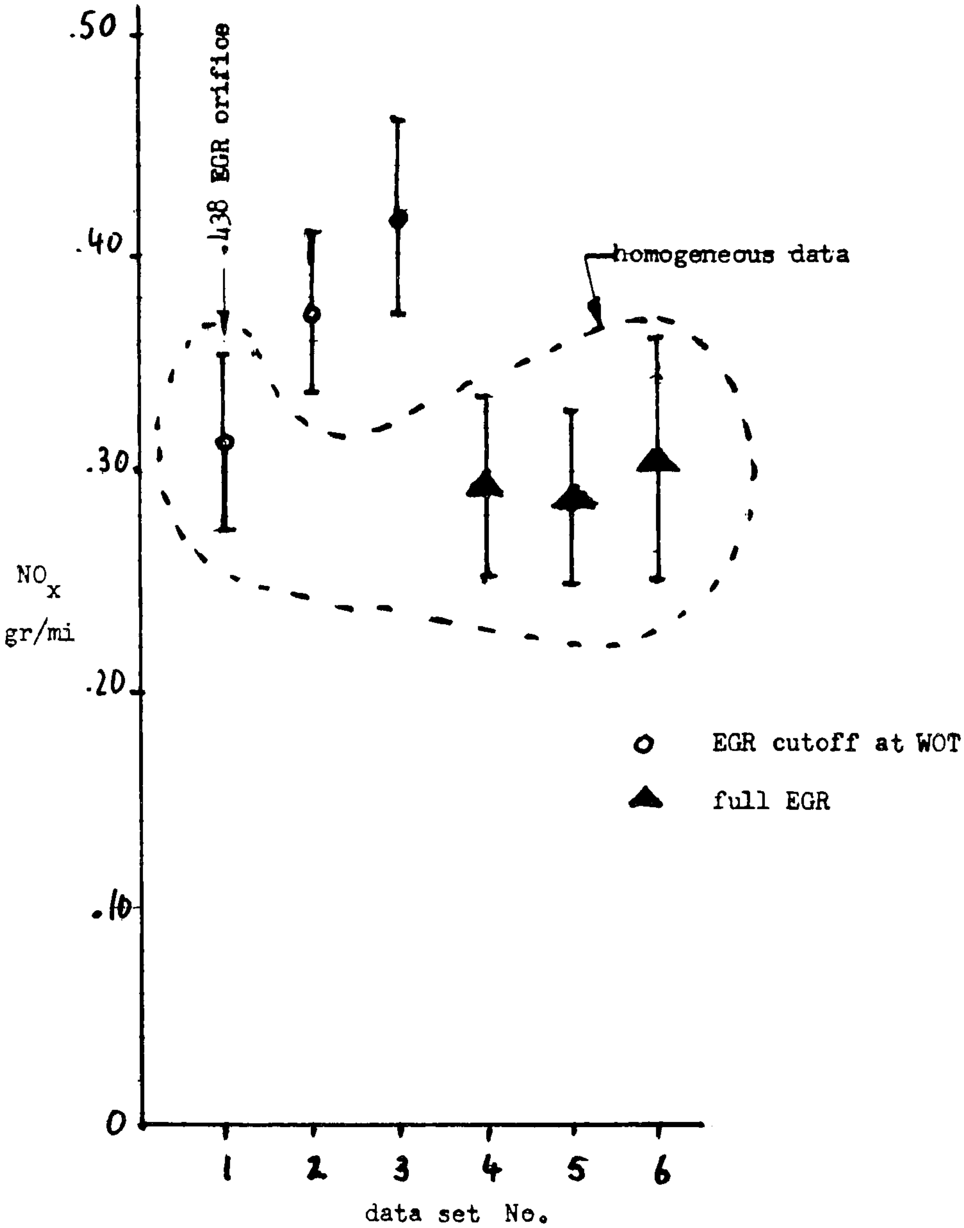


Figure 3

NO_x means and confidence ranges for six data sets

Data sets No. 1, 2, 6 Ford data

Data sets No. 3, 4, 5 EPA data

These levels combined with the earlier estimated standard deviations result in the following coefficients of variation:

HC coefficient of variation = (.0578/.371)100 ≈ 16% (based on data sets 1 + 3 + 4 + 5 + 6)

HC coefficient of variation = (.0578/.328)100 ≅ 17.5% (based on data sets 2 + 4 + 5 + 6)

CO coefficient of variation (.1347/.824)100 ≅ 16 % (based on data sets 3 + 4 + 5)

NOx coefficient of variation = (.0476/.298)100 ≈ 16% (based on data sets 1 + 4 + 5 + 6)

The coefficient of variation for all three emission gases is approximately the same, i.e., 16%.

In order to answer the question of whether the emission standards of 1976 are achieved we computed the percentage of data lower than the standards. Table 8 shows that 75% of the combined data sets (1 + 3 + 4 + 5 + 6) are lower than the HC standard of 0.41 gr/mi. If one excludes the runs with faulty catalyst and combines only data sets (2 + 4 + 5 + 6), then 92% of all data meet the HC standard. Whether these data demonstrate full compliance cannot be decided with confidence; a higher percentage would surely be desirable. Table 8 shows also that compliance with the CO-standard of 3.4 gr/mi is well achieved, and that compliance with the NOx standard of 0.40 gr/mi is within 98%.

We like to point out here that the given percentages are computed from the mean values and the standard deviations. Clearly, if the standard deviation of the HC data would have been lower, the HC standard could have been met. The standard deviation, as computed here, encompasses both the variations of the engine and those of the test apparatus and the test procedure including the operations of the test driver. If the variations of the test procedure and the test equipment play a significant part in the overall variation, a guess we cannot substantiate without

Exhaust gas Components	1976 National Standard gr/mi	Combined Data Sets	Number of Runs	Mean Value gr/mi	Standard Deviation gr/mi	Percentage of data below 1976 Standard
HC	0.41	1+3+4+5+6	20	0.37	0.058	75%
		2+4+5+6	16	0.33		92%
CO	3.4	3+4+5	13	0.82	0.135	100%
NO_x	0.40	1+4+5+6	20	0.30	0.048	98%

TABLE 8 Percentage of measured data below 1976 national standards

Data Set 1 FORD - faulty catalyst

2 FORD - new catalyst

3 EPA - maladjusted EGR

4 EPA - adjusted EGR

5 EPA - adjusted accelerator linkage

6 FORD

further experimentation, and if one could succeed in lowering these variations by using more accurate equipment or techniques, then compliance with the HC standard would not be unlikely.

Summary

A M-151 (jeep) vehicle equipped with a low-emission stratified-charge Ford engine (FCP) was emission-tested at both the Ford Company and EPA using a chassis dynamometer in combination with the 1976 CVS-CH testing and sampling technique. A total of 25 test runs structured into 6 test series were made with 2 to 5 runs per test series. With the exceptions of a correction of the EGR system and of the accelerator linkage, and the replacement of a faulty catalyst, no substantial changes were made during tests.

An analysis of the measured data revealed the following:

- The coefficient of variation of all three exhaust gas components is the same, i.e., 16 to 17 per cent.
- The replacement of the faulty PTX-6 palladium catalyst by a new PTX-6 platinum catalyst reduced HC and, especially, CO emission substantially.
- Elimination of EGR cutoff during accelerations reduced NO_x emissions substantially.
- Adjustment of the deceleration fuel cutoff linkage had little effect on emissions.
- For each gas component, a number of homogeneous data were selected as representative of the emission performance of the tested engine-vehicle combination. Of these data sets, 100% achieved compliance with the 1976 CO standard, 98% with the NO_x standard, and 92% with the HC standard if the runs with faulty catalyst were excluded. If they were included, then only 75% of all HC data achieved compliance with the HC standard.

REFERENCES

1 Ford Motor Comp, Special Military Vehicle Operations, Eng. Progress Rpt. No. 14 Project: Ford Combustion Process, July 1971

2 Ford Motor Comp, Special Military Vehicle Operations, Eng. Progress Rpt. No. 15 Project: Ford Combustion Process, August 1971

3 Ford Motor Comp, Special Military Vehicles Operation, Eng. Progress Rpt. No. 16 Project: Ford Combustion Process, September 1971

4 Computer printout of EPA, private communication

APPENDIX C - REGRESSION ANALYSIS OF EMISSIONS DATA

H. T. McAdams

1. Regression Analysis of Emissions Tests of a Stratified Charge Engine

A number of variables can affect emissions from a stratified charge engine. These include ambient environmental factors, such as barometric pressure, specific humidity and temperature, as well as such other variables as dynamometer inertia and road load horsepower. To estimate the effect of these variables on emissions, regression analysis was applied to 14 tests conducted on an FCP engine installed in a JEEP. The results of these tests, together with the associated levels of the ambient input variables, are tabulated in Table 1.

TABLE I

ANALYSIS OF 14 TESTS OF A STRATIFIED-CHARGE ENGINE

X_1 Bar.Press (in.Hg.)	X_2 Spec.Hum. (Lbs H_2O/ Lbs.Dry Air)	X_3 Temp. (°F)	X_4 Inertia (lbs)	X_5 Road Load Horsepower (HP)	Y_1 Hydrocarbons (Gms/Mi.)	Y_2 CO (Gms/Mi.)	Y_3 NOX (Gms./Mi)
29.3600	0.0105	81.0000	2750.0000	8.8000	0.4100	0.9800	0.4500
29.4600	0.0095	72.0000	2750.0000	8.5000	0.3700	0.7200	0.3900
29.4200	0.0115	76.0000	2750.0000	8.5000	0.4600	0.8100	0.4100
29.3800	0.0140	76.0000	2750.0000	8.5000	0.4600	0.8200	0.3700
29.3700	0.0143	75.0000	2750.0000	8.5000	0.3300	0.8100	0.4700
29.3200	0.0133	74.0000	2750.0000	8.5000	0.3700	0.9200	0.2100
29.2500	0.0147	81.0000	2750.0000	8.5000	0.4700	1.1700	0.3400
29.2500	0.0148	83.0000	3000.0000	8.8000	0.3900	0.9200	0.3200
29.3000	0.0125	76.0000	3000.0000	8.8000	0.3200	0.8300	0.3100
29.2800	0.0114	81.0000	3000.0000	8.8000	0.2700	0.5800	0.2800
29.1800	0.0116	74.0000	3000.0000	8.8000	0.3000	0.5500	0.2900
29.0300	0.0101	70.0000	3000.0000	8.8000	0.3200	0.7800	0.2600
29.0000	0.0097	72.0000	3000.0000	8.8000	0.3100	1.2000	0.2500
28.9900	0.0099	70.0000	3000.0000	8.8000	0.4000	1.2400	0.3500

As an approach to studying the effect of ambient conditions on emissions, it can be supposed that the emissions Y, whether hydrocarbons (Y_1), CO (Y_2), or NOX (Y_3) can be expressed as a function of the variables X_1, X_2, X_3, X_4, and X_5.

$$Y = f(X_1, X_2, X_3, X_4, X_5)$$

The simplest relationship which can be hypothesized is a linear function

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5$$

where b_0, b_1, b_2, b_3, b_4 and b_5 are constants. Such a regression equation can be fitted to the data of Table I by the method of least squares and its statistical significance tested by analysis of variance.

The regression equations computed from the data of Table I are given below.

Hydrocarbons:

$$Y_1 = 8.93979 - 0.23550\,X_1 - 0.94665\,X_2 + 0.00590\,X_3 - 0.00038\,X_4 - 0.11858\,X_5$$

CO :

$$Y_2 = 56.83229 - 1.78644\,X_1 - 5.20022\,X_2 + 0.01833\,X_3 - 0.00150\,X_4 - 0.07912\,X_5$$

NOX:

$$Y_3 = 6.45692 + 0.18473\,X_1 + 2.83481\,X_2 - 0.00263\,X_3 - 0.00054\,X_4 + 0.35757\,X_5$$

For each of the regression equations, the inputted or observed values of emission were compared with the corresponding values computed from the regression equation. For example, see Table II, in which the observed and computed values of hydrocarbons are tabulated.

TABLE II

COMPARISON OF OBSERVED AND CALCULATED HYDROCARBON EMISSIONS

Y_1 Observed	Y_1 Calculated
0.41000	0.40983
0.37000	0.36969
0.46000	0.40082
0.46000	0.40788
0.33000	0.40404
0.37000	0.41087
0.47000	0.46733
0.39000	0.34889
0.32000	0.29799
0.27000	0.33324
0.30000	0.31530
0.32000	0.32843
0.31000	0.34768
0.40000	0.33805

By computing the difference between the observed and calculated quantities for each test, squaring these differences and summing over all tests, one obtains a measure of the residual scatter unaccounted for by a linear regression equation. This residual sum of squares can be considered as a measure of error, on the assumption that the regression equation, as formulated, extracts all systematic (non-random) trends in the data. An alternative assumption is that there is no relation between the observed emissions and the values of the ambient variables. Accordingly, one can compute the average value of the emissions for the 14 tests and compare the 14 observed emissions with this single average value. If the deviations from the average value are computed, squared and summed, one obtains a residual sum of squares which remains when only a constant (the mean) is removed from the data. This residual sum of squares can not be smaller than the sum of squares of residuals obtained in the complete regression equation. The difference in magnitude of these two residual sums of squares is a measure of the variability accounted for by adjustment for linear trends with barometric pressure, specific humidity, temperature, load and horsepower. Whether this difference is large enough to be considered statistically significant can be adjudged by analysis-of-variance techniques applied to the regression results.

Tables III, IV and V summarize the results of analysis of variance for HC, CO and NOX respectively. In each table there is displayed the sum of squares due to b_0, b_1, b_2, b_3, b_4 and b_5 and the sum of squares due to the single constant b_0', the mean for all the data.* The salient feature of each of these tables is the F-ratio displayed in the right-most column. In order to argue that the effect of the input variables is statistically significant at the 5% level of significance the tabulated F-ratio should exceed the critical value

$$F_8^5\ (0.05) = 3.69$$

* Note that b_0' is different from b_0 in the regression equation. Though the two quantities play a similar role, that of a constant in an empirical equation, they will be equal only if the values of the variables X_1, X_2, X_3, X_4 and X_5 constitute an orthogonal set of vectors in the matrix formulation of the least-squares normal equations.

TABLE III

ANALYSIS OF VARIANCE FOR REGRESSION OF HC EMISSIONS FOR A STRATIFIED CHARGE ENGINE

Source of Variance	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Total	14	1.971		
Due to $b_0, b_1, b_2, b_3, b_4, b_5$	6	1.946		
Due to b_0	1	1.917		
Due to b_1, b_2, b_3, b_4, b_5	5	.029	.0058	1.856
Residual	8	.025	.0031	

$F_8^5 \ (.95) \doteq 3.69$

ANALYSIS OF VARIANCE FOR REGRESSION ANALYSIS OF CO EMISSIONS
FOR A STRATIFIED CHARGE ENGINE

Source of Variance	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Total	14	11.488		
Due to $b_0, b_1, b_2, b_3, b_4, b_5$	6	11.246		
Due to b_0'	1	10.912		
Due to b_1, b_2, b_3, b_4, b_5	5	.334	.0668	2.208
Residual	8	.242	.0302	

F_8^5 (.95) = 3.69

TABLE V

ANALYSIS OF VARIANCE FOR REGRESSION ANALYSIS OF NOX EMISSIONS FOR A STRATIFIED CHARGE ENGINE

Source of Variance	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Total	14	1.654		
Due to $b_0, b_1, b_2, b_3, b_4, b_5$	6	1.612		
Due to b_0'	1	1.578		
Due to b_1, b_2, b_3, b_4, b_5	5	.034	.0068	1.295
Residual	8	.042	.0053	

F_8^5 (.95) = 3.69

Values of F smaller than 3.69 could occur relatively often, even if there were no effect of the input variables, simply as the result of random errors. A value as large or larger than 3.69, however, would occur by chance no more than 5% of the time. Therefore, if the critical value is exceeded by the F-ratio computed from the data, it can be argued that the computed regression equation is <u>not</u> the result of chance but reflects an actual relationship between the dependent and the independent variables. It will be noted that for neither HC, CO nor NOX did the computed F-ratio exceed the critical value. Therefore, it must be concluded that the data from the 14 tests do not establish any linear dependence of emissions on atmospheric pressure, specific humidity. temperature, inertia and horsepower. Since the combined effects of these five variables failed to reach the desired significance level, it can further be concluded that any single variable would also fail to show a significant influence on emissions.

2. Regression Analysis of Emissions Tests of Conventional Engines

Emissions tests are available on 54 automobiles representing various engine sizes, vehicle weights and mileages. It was desired to determine if emissions are related to these factors, as judged by regression analyses. For purposes of the analysis, the engines were segregated into two groups, 8-cylinder and 6-cylinder. Because of some questionable aspects of the data for one of the vehicles, only 53 of the 54 items were actually analyzed.

Input data for the 8-cylinder engines, 45 in number, are given in Table VI. Regression equations for these data took the form

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3$$

where X_1 - mileage (mi.)
X_2 = weight (lbs)
X_3 = displacement (cu.in.)

and Y is either HC, CO or NOX in grams/mile.

The regression equations as derived from the data of Table VI are given below.

HC:

$$Y_1 = 1.05587 - 0.00000\ X_1 - 0.00004\ X_2 + 0.00008\ X_3$$

CO:

$$Y_2 - 5.27079 + 0.00002\ X_1 - 0.00106\ X_2 + 0.01650\ X_3$$

NOX :

$$Y_3 \quad 0.49340 - 0.00000\ X_1 + 0.00025\ X_2 \quad 0.00045\ X_3$$

Analysis of variance for these three equations is given in Tables VII, VIII, and IX respectively.

TABLE VI

EMISSIONS DATA FOR 8-CYLINDER ENGINES

X_1 Mileage (Mi.)	X_2 Weight (lbs)	X_3 Displacement (cu.in.)	Y_1 HC (gms/mi.)	Y_2 CO (gms/mi.)	Y_3 NOX (gms/mi.)
14173	5000	475	0.9800	8.4400	1.0600
23801	4500	475	0.7930	6.6300	1.5300
24930	4500	360	0.8250	3.9200	1.1300
15427	3500	360	1.0200	8.1100	1.0300
15598	3500	360	1.2200	11.8700	1.1400
18923	4000	410	0.6670	8.1400	0.8200
13625	4000	360	0.7960	4.7600	1.1400
8039	4000	360	0.7800	6.4800	1.3100
22772	3500	290	0.7840	7.7400	1.3200
23957	4500	410	0.7210	2.0500	1.5500
9803	4000	360	0.8790	4.9100	1.5400
13664	3500	290	1.2300	5.0500	1.2800
20760	4000	360	0.7500	12.0400	1.1800
22188	4000	410	0.9930	11.1700	1.5800
14700	3500	290	0.6840	3.1900	1.0200
31888	5000	475	0.7310	5.9800	2.5000
9943	4000	360	0.8740	9.7300	1.5500
9797	3500	360	1.0600	9.7400	0.8630
30064	5000	410	0.7900	5.7000	1.5200
20035	4500	475	0.7960	6.0700	1.6700
17756	4000	360	0.8010	5.2200	1.2700
11089	5000	410	1.3200	13.2500	2.4800
29118	4500	475	1.0900	16.9900	0.5910
24963	4500	475	0.6840	8.5500	1.1300
18686	3000	360	1.0500	9.8800	1.2300
27205	4000	360	0.8920	8.6700	1.0500
28565	4500	475	0.8580	8.8700	1.2500
6078	4000	410	0.7340	4.2100	1.4800
10229	4000	360	0.7510	5.6700	1.7800
11312	5000	410	0.7750	4.9300	1.1600
32840	5500	475	0.8230	10.4700	1.2800
13866	3500	290	0.8520	9.6500	0.9570
15485	5500	475	0.5730	10.0200	0.7420
9351	3500	360	0.8880	11.8900	1.0500
20831	4000	410	0.8320	8.1600	1.2000
20204	4000	360	0.8610	5.9600	1.1800
20037	4500	290	0.8910	7.4900	1.3700
31419	4500	475	0.6910	8.7100	1.3700
13032	4500	475	0.8680	5.7300	0.9720
20206	3500	290	0.7820	4.4500	0.9520
8624	4000	360	0.8240	7.8600	1.1400
8991	4500	410	0.9130	7.4500	1.7900
16698	3500	360	0.5910	2.8600	0.9570
34397	4500	290	0.7570	6.1200	1.2700
30749	5000	475	0.8370	5.8300	1.4700

* Displacement is the median value of a class interval, the engines being identified only according to a range of displacements.

TABLE VII

ANALYSIS OF VARIANCE FOR REGRESSION ANALYSIS OF HC EMISSIONS FOR 8-CYLINDER ENGINES

Source of Variance	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Total	45	33.704		
Due to b_0, b_1, b_2, b_3	4	32.677		
Due to b_0'	1	32.616		
Due to b_1, b_2, b_3	3	.061	.020	0.800
Residual	41	1.027	.025	

F^{3}_{41} (.05) = 2.83

TABLE VIII

ANALYSIS OF VARIANCE FOR REGRESSION ANALYSIS OF CO EMISSIONS FOR 8-CYLINDER ENGINES

Source of Variance	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Total	45	2974.730		
Due to b_0, b_1, b_2, b_3	4	2605.221		
Due to b_0	1	2578.112		
Due to b_1, b_2, b_3	3	27.109	9.036	1.003
Residual	41	369.509	9.012	

$F^{3}_{47}(.05) = 2.83$

TABLE IX

ANALYSIS OF VARIANCE FOR REGRESSION ANALYSIS OF NOX EMISSIONS FOR 8-CYLINDER ENGINES

Source of Variance	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Total	45	80.480		
Due to b_0, b_1, b_2, b_3	4	75.008		
Due to b_0	1	74.379		
Due to b_1, b_2, b_3	3	.629	.210	1.57
Residual	41	5.472	.133	

F^{3}_{41} (.05) = 2.83

It is seen, by reference to these tables, that statistical significance at the 0.05 level is not achieved by either HC, CO or NOX even when the emissions are adjusted for all three variables X_1, X_2 and X_3. Therefore, it must be concluded that no one of these variables would show a satistically significant effect in the cases analyzed.

A word of explanation is in order concerning the displacement variable, X_3. Exact displacement for each engine was unknown. Rather, each vehicle was assigned to a class, as shown below.

Class	Range of Engine Displacement (cu.in.)
1	Less than 260
2	261 to 320
3	321 to 380
4	381 to 439
5	Greater than 440

In applying the regression analysis, mean or representative values were assumed for the five classes. These were 290, 360, 410, and 475 for classes 2, 3, 4 and 5, respectively. No assumption was necessary for Class 1, since this included only six-cylinder engines, which were handled separately from eight-cylinder engines. By coding the displacement values in this way, it was possible to avoid identifying actual makes or manufacturers of the vehicles tested.

A separate analysis was performed for the 6-cylinder engines, eight in number. The input data for these engines is tabulated in Table X. Inasmuch as all of the 6-cylinder engines belonged to the same displacement class, displacement was not a variable for these data. Therefore, the emissions were expressed as functions of only X_1, mileage and X_2, weight.

TABLE X

EMISSIONS FOR 6-CYLINDER ENGINES

X_1 mileage	X_2 Weight	HO (gms/mi.)	CO (gms/mi.)	NOX (gms/mi.)
18618	3500	0.524	4.18	1.72
12320	3500	0.861	16.51	1.21
12769	3000	0.478	3.81	1.13
4106	2500	0.533	3.95	1.04
4395	3500	0.437	4.79	1.60
10923	2500	0.731	3.89	1.83
16183	3000	0.665	3.90	0.72
15124	3500	0.926	13.66	2.07

The regression equations, as derived from the data of Table X, follow.

Hydrocarbons:

$$Y_1 \quad 0.44245 + 0.00001\, X_1 + 0.00002\, X_2$$

CO:

$$Y_2 = -13.23511 - 0.00002\, X_1 + 0.00652\, X_2$$

NOX:

$$Y_3 = 0.39274 + 0.00000\, X_1 + 0.00033\, X_2$$

Analysis of variance for these three equations are given in Tables XI, XII and XIII, respectively.

It is seen, by reference to these tables, that statistical significance at the 5% level is not achieved.

In summary, it is concluded that no statistical significance is indicated for the regression equations computed in this Appendix. The effects of the variables in question either are non-existent or are too small, relative to error magnitudes, to be detectable within the limited sample of available data.

TABLE XI

ANALYSIS OF VARIANCE FOR REGRESSION ANALYSIS OF HC EMISSIONS FOR 6-CYLINDER ENGINES

Source of Variance	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Total	8	3.554		
Due to b_0, b_1, b_2	3	3.354		
Due to b_0'	1	3.322		
Due to b_1, b_2	2	.032	.016	0.400
Residual	5	.200	.040	

F_5^2 (.05) = 5.79

TABLE XII

ANALYSIS OF VARIANCE FOR REGRESSION ANALYSIS OF CO EMISSIONS FOR 6-CYLINDER ENGINES

Source of Variance	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Total	8	560.052		
Due to b_0, b_1, b_2	3	430.472		
Due to b_0'	1	373.874		
Due to b_1, b_2	2	56.598	28.299	1.092
Residual	5	129.580	25.916	

$F_5^2\ (.05) = 5.79$

TABLE XIII

ANALYSIS OF VARIANCE FOR REGRESSION ANALYSIS OF NOX EMISSIONS FOR 6-CYLINDER ENGINES

Source of Variance	Degrees of Freedom	Sum of Squares	Mean Squares	F Ratio
Total	8	17.493		
Due to b_0, b_1, b_2	3	16.165		
Due to b_0'	1	16.018		
Due to b_1, b_2	2	.147	.074	0.278
Residual	5	1.328	.266	

$F^2_5\ (.05) = 5.79$

APPENDIX D ANALYSIS OF EMISSION TEST DATA OF VARIOUS GROUPS OF AUTOMOTIVE VEHICLES

D. J Schuring

INTRODUCTION

All experimental data are subject to fluctuations ("errors" or "noise") caused by erratic, usually small, changes of the experimental equipment during measurements. In emission testing, the fluctuations are often large, that is, not much smaller than the differences ("signals") we wish to study, so that statistical techniques must be employed to separate the signals from the noise.

Emission fluctuations originate from many sources -- fuel injection, spark plugs, air temperature, humidity, air-fuel mixture, compression ratio, combustion process, valves, linkages, collecting and diagnostic equipment, and the human operator -- to name a few. The fluctuations issuing from these and many other sources are by necessity confounded; they can hardly be separated except for the hypothetical case of all but one of the errors being negligibly small. The fluctuations of an engine (taken as a unit source of error), for instance, could be isolated only if the error of the measuring equipment would be negligibly small, if each measurement would follow exactly the same procedure, and if the initial engine and vehicle conditions before each repetition, and the temperature, humidity, and pressure of the air, as well as all other external variables, would remain constant. Only under these (hypothetical) conditions would the tested engine reveal its inherent noise level.

In reality, of course, irregularities of the measuring equipment, deviations from the prescribed test procedure, variations of the atmospheric state, and changes of the initial vehicle conditions are inevitable. Hence, each repetition yields a different result. The fluctuations of observations performed repeatedly on the same car with the same equipment under identically controlled test conditions will be denoted by the term "test error".

The test error masks the emission differences we are interested in -- differences between cars at the end of the production line, between performances of the same car at various stages of its life, between cars of different make, between cars with and without emission control, and the like. The question then arises of how many cars have to be tested in order to establish the desired information with confidence. The following study attempts to shed some light on this problem. It tries a first cut at the relation between the test error, which is present in all measurements, and the emission differences between cars. Unfortunately, the available pertinent information is far from complete. The data used here are gleaned from some EPA reports, some company presentations, and from communications of automobile manufacturers. Since the latter were confidential, they could not be fully utilized. Furthermore, the size of available emission samples was usually small -- frequently not larger than five to ten observations per sample. Under these circumstances, lacking statistical evidence had to be filled in by engineering judgment. In addition, statistical data for very low emission engines are practically non-existent because these engines are still under development. Notwithstanding all these restrictions, we felt that any, however preliminary, information on the fluctuations of emission measurements would be helpful in planning tests for the stratified-combustion-process engines. Hence, this study.

EMISSION DISTRIBUTIONS

The test error, that is, the distribution of all observations performed repeatedly on the same car with the same equipment under identically controlled test conditions, can be characterized statistically in various ways. Since emission distributions are usually not normal, a characterization by the mean and the standard deviation is often not sufficient. A more complete picture evolves from the cumulative relative frequency distribution: data are grouped into equidistant classes and their relative frequency, i. e., the relative number of data in each class, is associated with the respective upper class level. Plotted on probability paper (on which the ordinate scale is graduated according to the area under a normal distribution), the cumulative data, if essentially normally distributed, approximate a straight line. The mean value appears then at the cumulative relative frequency of 50%, the standard deviation as difference of the emission values at the frequencies of 84% and 50% (or 16% and 50%).

Emission data are not normally distributed, however, as indicated before. They are skewed because emissions cannot surpass a certain minimum value (ultimately, the zero value). If the mean value of an emission distribution is large (as is true for an uncontrolled engine), the skew may be slight and can perhaps be neglected. Then, a normal distribution would fit the data best. If the mean value is close to the zero point, however, as realized for an emission-controlled engine, the emission data are compressed towards smaller values, and expanded towards larger values. Skewed distributions of this kind can often be turned into normal distributions by converting the measured data into their log values before processing them statistically. The transformation is most easily achieved by plotting the relative accumulative frequencies on probability paper with a logarithmically graduated abscissa scale. If the log values of the emission data are normally distributed, they will approximate a straight line on this paper. The emission value at the cumulative frequency of 50% signifies the median rather than the mean (which is somewhat larger than the median). Also, the standard deviation cannot be directly obtained from this log plot. Instead, the percentage of emission data smaller (or larger) than any given emission value can be predicted with ease.

As an example, we present HC and CO emission data of a large number of 1970 General Motors cars. The cars had been in actual customer use and were tested by the California Air Resources Board using the FTP hot-cycle procedure. Figures 1 and 2 show the data as presented by GM. We subdivided the data into 12 classes (each 0.50 gr/mi wide), counted the number of observations within each class, computed their relative accumulated frequencies, plotted them at the upper level of each class on log-probability paper, as shown in Figure 3 for HC and Figure 4 for CO, and faired each plot by a straight line signifying a logarithmic distribution of emission data. The data at the ends are somewhat erratic, which must be attributed to the small number of observations at high emission levels, see Figures 1 and 2.

Since these and other data (see Figures 5, 6, 7) indicate strongly that emissions data generally conform to the normal log-distribution, we plotted all emission distributions on log-probability paper and faired them by straight lines even if the scatter was large. The scatter was always attributed to the small number of observations rather than to a significant deviation from the log-distribution. We would like to note that the emission data in Figures 1 and 3 were obtained from many vehicles instead of just one; they exhibit the combined effect of the test error and the variance

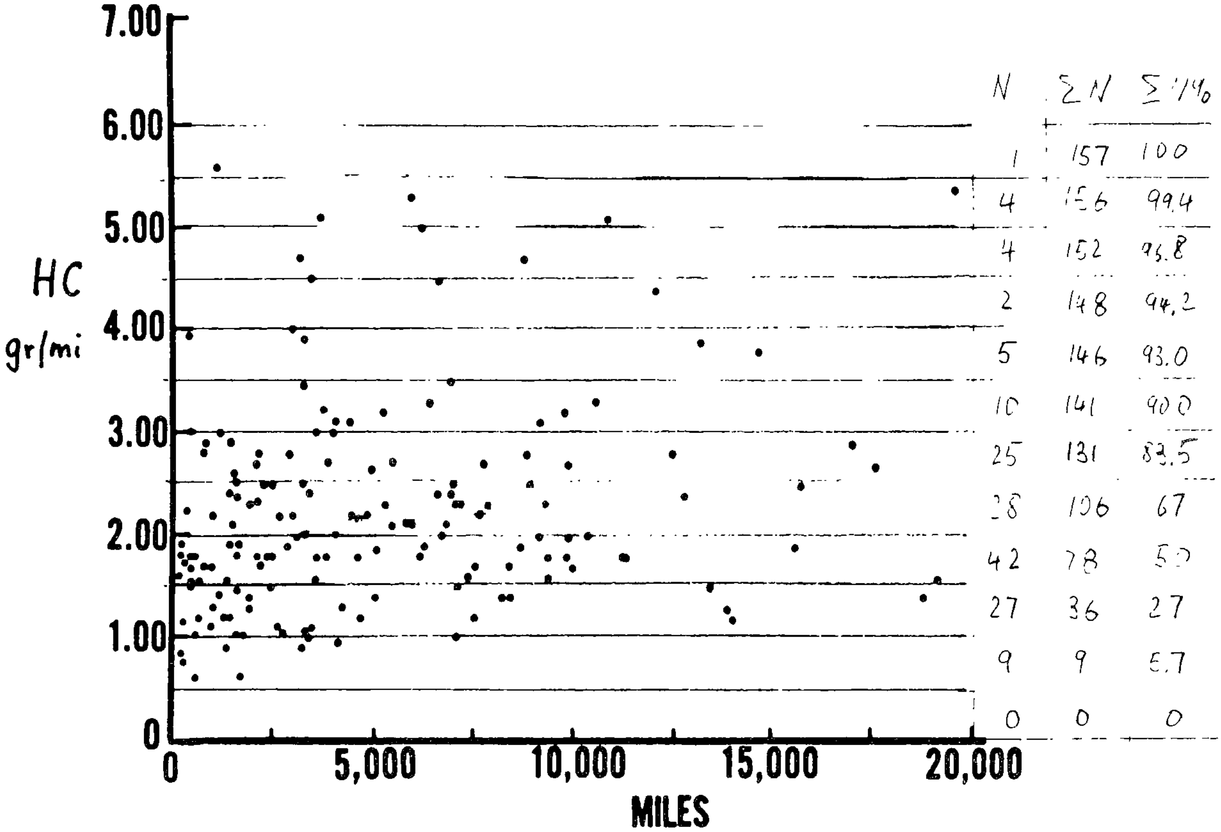


Figure 1 - HC Emissions of GM Cars, Model 1970, Tested by California Air Resources Board (FTP Hot-Cycle) Adapted from Reference 1

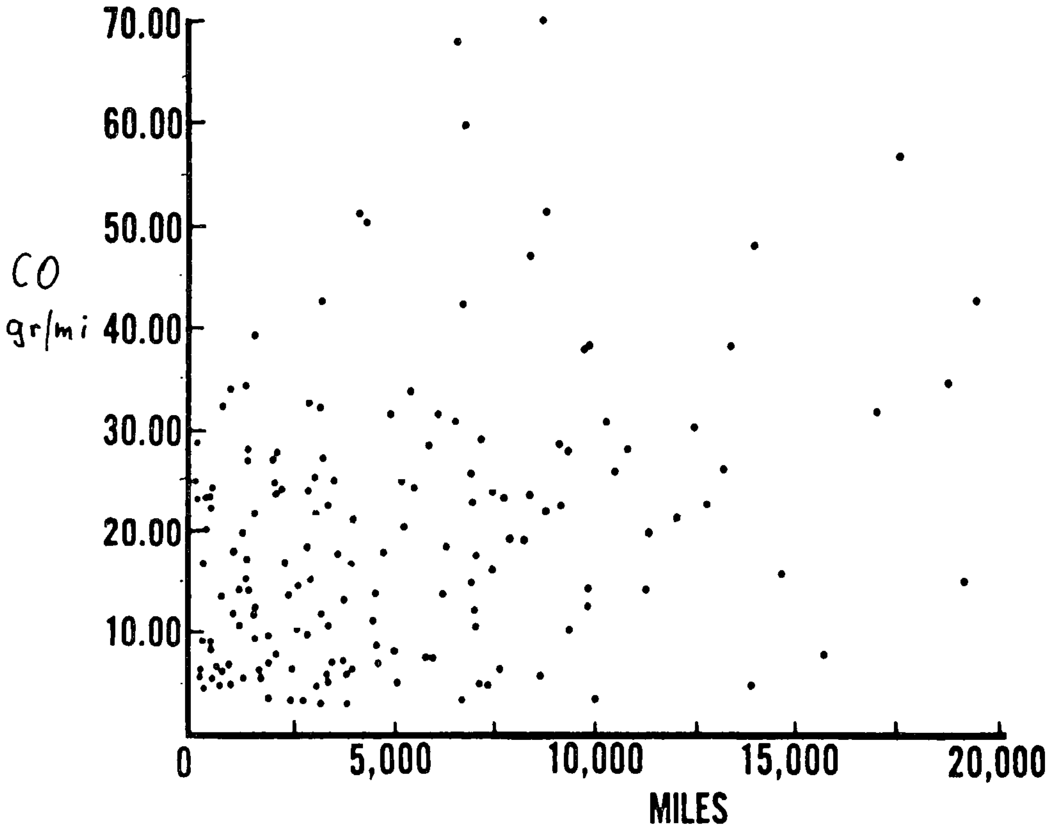


Figure 2 - CO Emissions of GM Cars, Model 1970, Tested by California Air Resources Board (FTP Hot-Cycle) Adapted from Reference 1

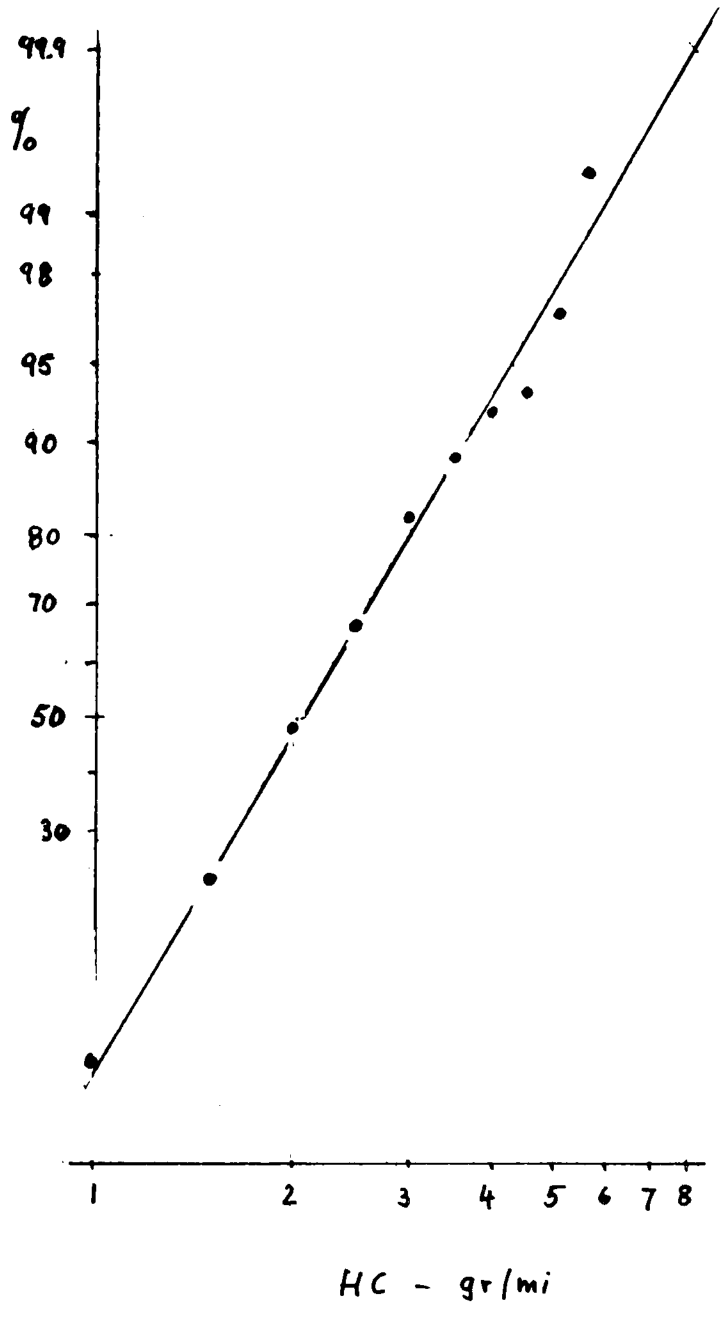


Figure 3 - Cumulative Relative Frequency Distribution of HC Emissions of GM Cars, Model 1970, Tested by California Air Resources Board (FTP Hot-Cycle) - Reference 1

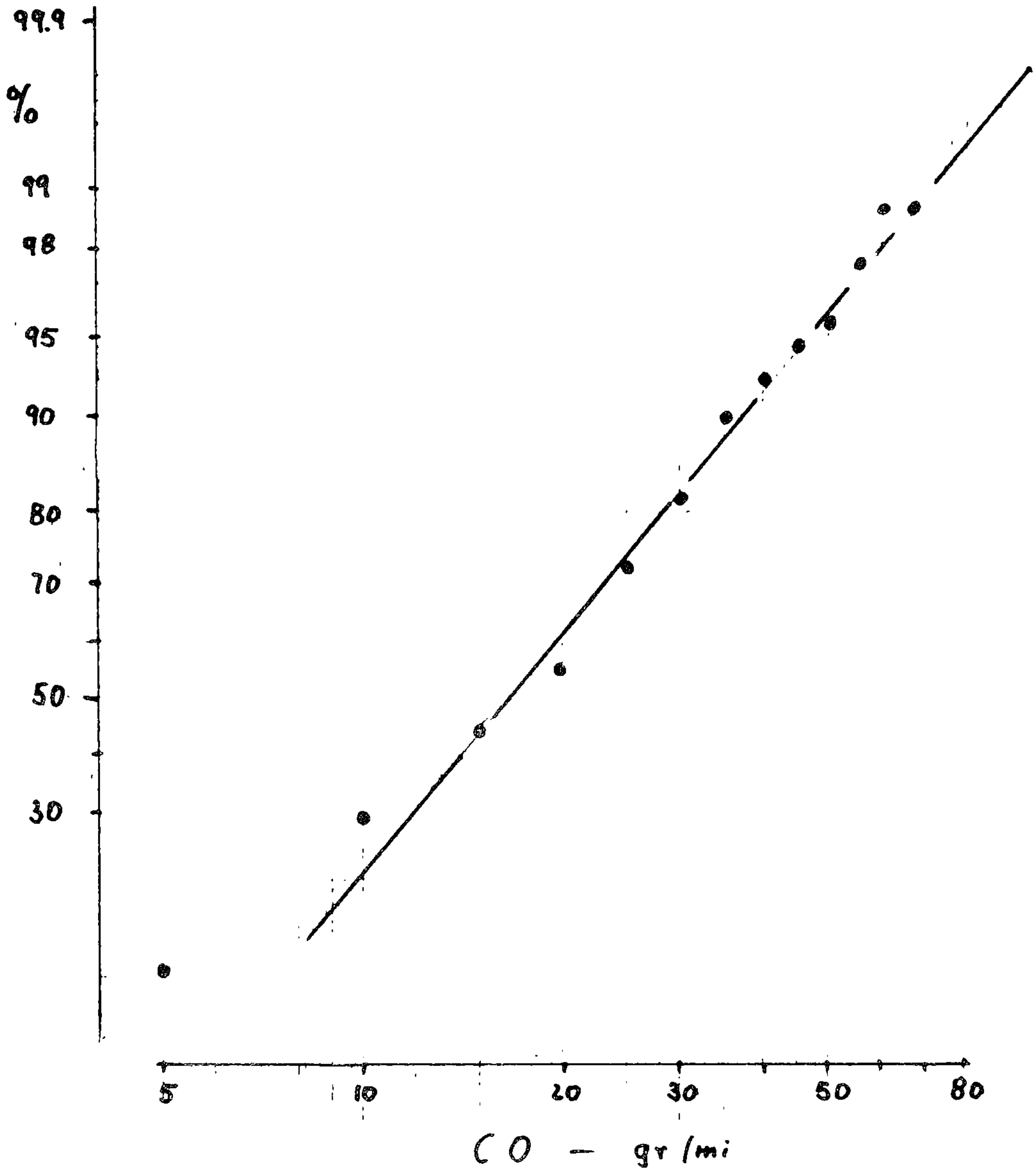


Figure 4 - Cumulative Relative Frequency Distribution of CO Emissions of GM Cars, Model 1970, Tested by California Air Resources Board (FTP Hot-Cycle) - Reference 1

between vehicles. Unfortunately, no single vehicle measurements with a large number of repetitions were available. We will assume, however, that all emission data, whether obtained from one engine or a large number of engines, are distributed logarithmically.

In the following, a number of emission measurements gleaned from papers, company reports, EPA evaluations, and NRC* presentations are plotted on log-probability paper. In instances where only the mean and the standard deviation were available instead of the raw data, the following transformation was made.

If the mean, $\bar{x}$, and the standard deviation, s, are given of a set of observations, x_i, whose log-distribution we know (or suspect) to be a normal one, we can compute three characteristical values: the median, x_{50}, (which is defined as the upper bound of 50% of all observations), the value x_{84}, bounding 84% of all observations, and the value x_{16}, 16%. The formulas for computing these three values from $\bar{x}$ and s are given below. The derivation is presented in Appendix G.

median
$$x_{50} = \frac{\bar{x}^2}{\sqrt{\bar{x}^2 + s^2}},$$

84% limit**
$$x_{84} = x_{50} \exp\sqrt{\ln\left[\left(\frac{s}{\bar{x}}\right)^2 + 1\right]}$$

16% limit
$$x_{16} = \frac{x_{50}}{\exp\sqrt{\ln\left[\left(\frac{s}{\bar{x}}\right)^2 + 1\right]}} = \frac{x_{50}^2}{x_{84}}$$

---

* National Research Council - National Academy of Engineering - Committee on Motor Vehicle Emissions

** $\exp \Gamma \equiv e^{\Gamma}$

These three data fix three points on log-probability paper at 50%, 84%, and 16% of the cumulative relative distribution scale. A (straight) line through these points establishes the log-distribution, from which any data of interest can be determined easily.

DATA

(1) In a Progress Report to EPA,[1] the General Motors Corporation published HC emissions and CO emissions of a large number of 1970 GM cars tested by the California Air Resources Board using the FTP hot-cycle test procedure. In this report the emission data were displayed as function of mileage in two graphs, Figures 1 and 2. The scatter is very large; hence, an influence of mileage on emission is unlikely, and we considered all data as drawn from a homogeneous population of cars.

The number of observations is approximately 160. A visual inspection of the distribution of data points suggests that repetitive measurements had not been made so that each point was assumed to represent the emission of one car measured only one time at given mileage.

A log-distribution plot of all CO and HC data is presented in Figures 3 and 4.

(2) Fifty-four fleet cars of different makes and mileage were tested by EPA[2] using the CVS-CH cycle (41 minutes). The identity of the vehicles was not disclosed. An analysis of variance (discussed elsewhere) did not disprove the assumption of all cars representing a sample of a homogeneous population. Hence, all HC, CO, and NOx data were lumped together and their distribution plotted, Figures 5, 6, and 7.

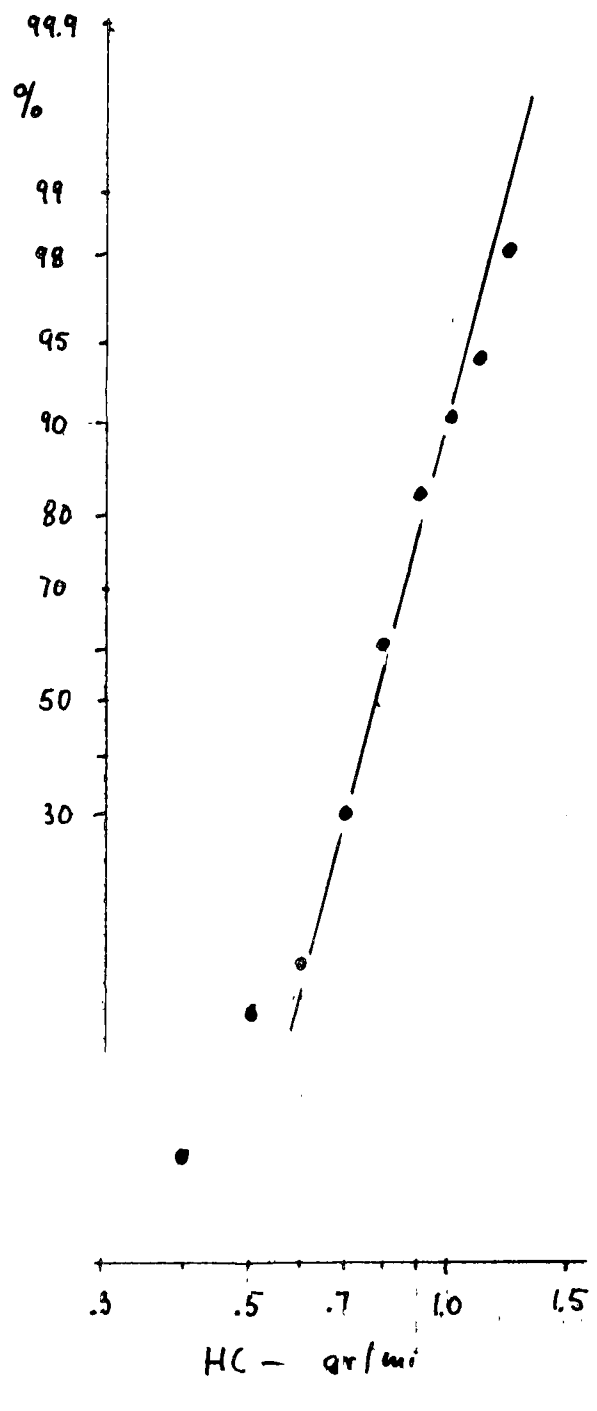


Figure 5 - Cumulative Relative Frequency Distribution of HC Emissions of 54 Fleet Cars Tested by EPA (CVS-CH Cycle)-Reference 2

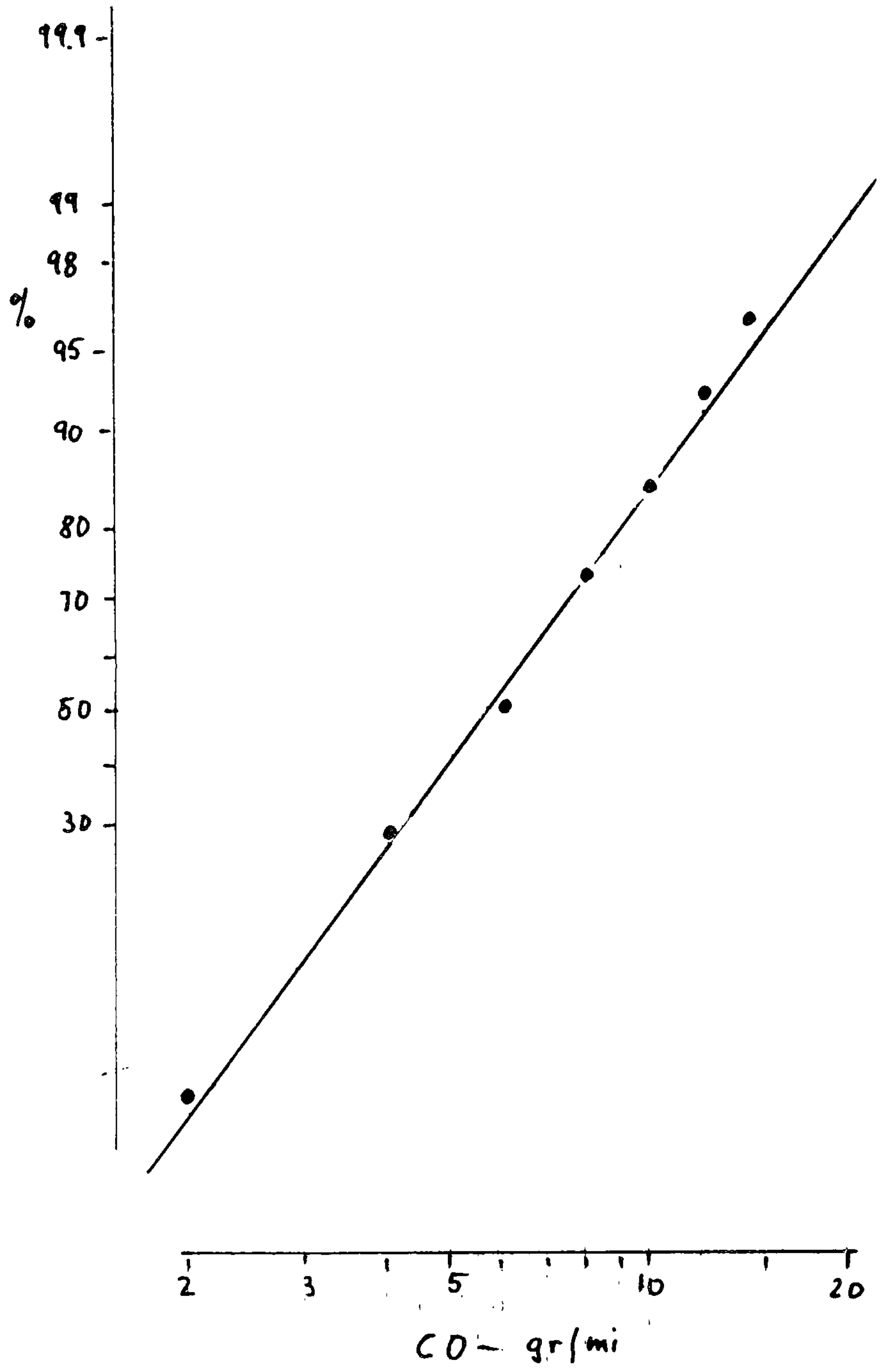


Figure 6 Cumulative Relative Frequency Distribution of CO Emissions of 54 Fleet Cars Tested by EPA (CVS-CH Cycle) - Reference 2

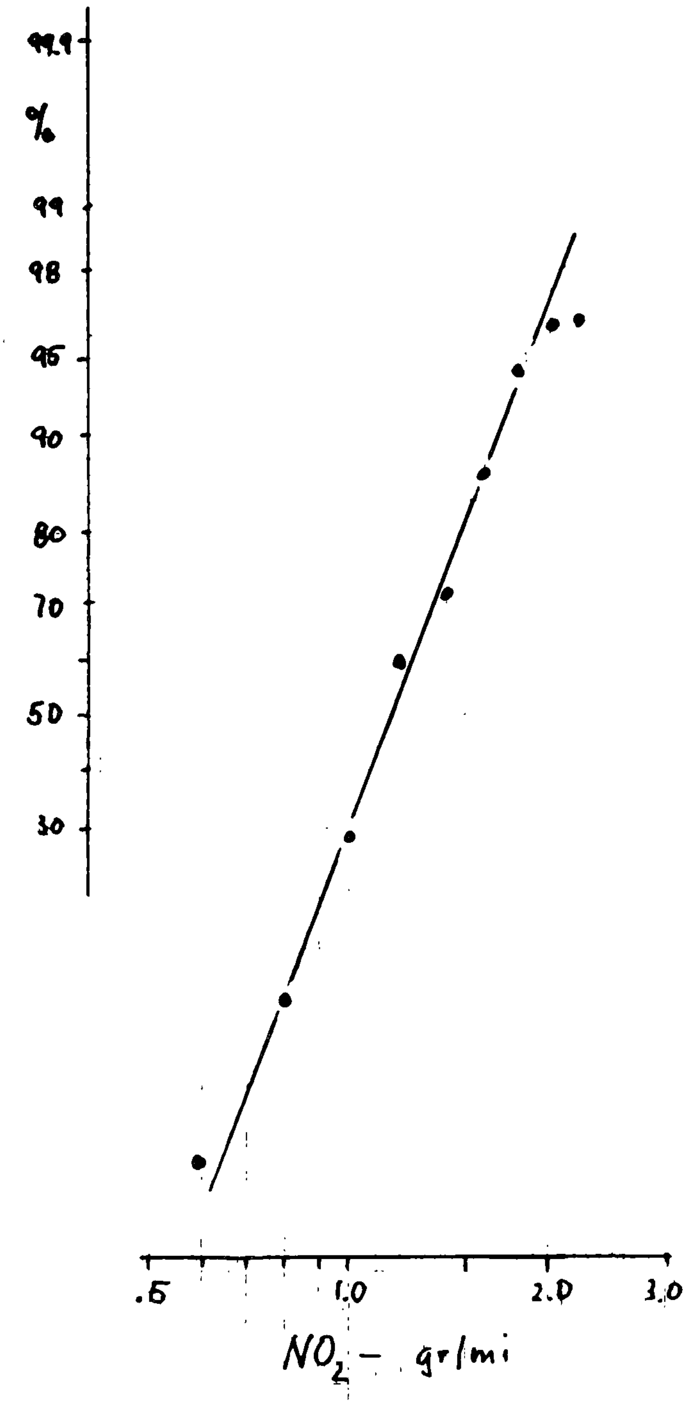


Figure 7 - Cumulative Relative Frequency Distribution of NO_2 Emissions of 54 Fleet Cars Tested by EPA (CVS-CH Cycle) - Reference 2

(3) EPA[3] reported on emissions of two new (less than 60 miles) M-151 Jeeps, production 1971, tested according to the 1972 CVS-C test procedure (23 minutes). The raw data are shown in Table 1. A t-test and a F-test revealed no significant differences between the two data sets, neither in the variance nor in the means. Consequently, all data were combined and their distribution plotted, Figure 8.

(4) A company whose identity cannot be disclosed[4] established limits of mean and standard deviations that, in the company's judgment, would have to be passed by all cars if they were to comply with the 1976 emission standards. Table 2 shows recommended mean and standard deviations together with the 1976 standards. From them, we computed the median, x_{50}; and x_{16} and x_{84} of the log-distribution. Figure 9 shows the results on log-probability paper.

(5) EPA measured the exhaust emissions of nine Rebels and nine Falcons,[5] all models 1970, using the CVS-C sampling technique in combination with the 23-minute driving schedule as specified for 1972. Tables 3 and 4 list the results. Figures 10 and 11 picture their log-distribution.

(6) Another company whose identity cannot be disclosed[6] communicated emission results obtained from two unidentified cars by the CVS-C test procedure (23 minutes). Both cars were equipped with experimental emission control systems. Emissions were measured at various mileages, Table 5, but no trend was discernible. Consequently, the six observations per emission constituent and car were considered repetitions. The log plot of their distribution is shown in Figure 12. The same company revealed emission data of a large group of cars of the same make tested by the hot 7-mode cycle. The results of their evaluation are plotted in Figure 13.

Table 1

1972 Federal Emission Test of 1971 Army M-151 Jeep
Performed by EPA
All Numbers in Grams per Mile
(Reference 3)

	HC	CO	NOx
Vehicle No. 1	5.0 5.5 3.9 7.5 4.4 5.0 3.2	150+ 128 101 150+ 110 109 87	0.6 1.9 2.1 1.2 2.1 2.1 1.3
Vehicle No. 2	6.2 4.8 7.3 5.0 5.9	103 115 137 103 102	1.8 3.0 2.3 1.6 1.8

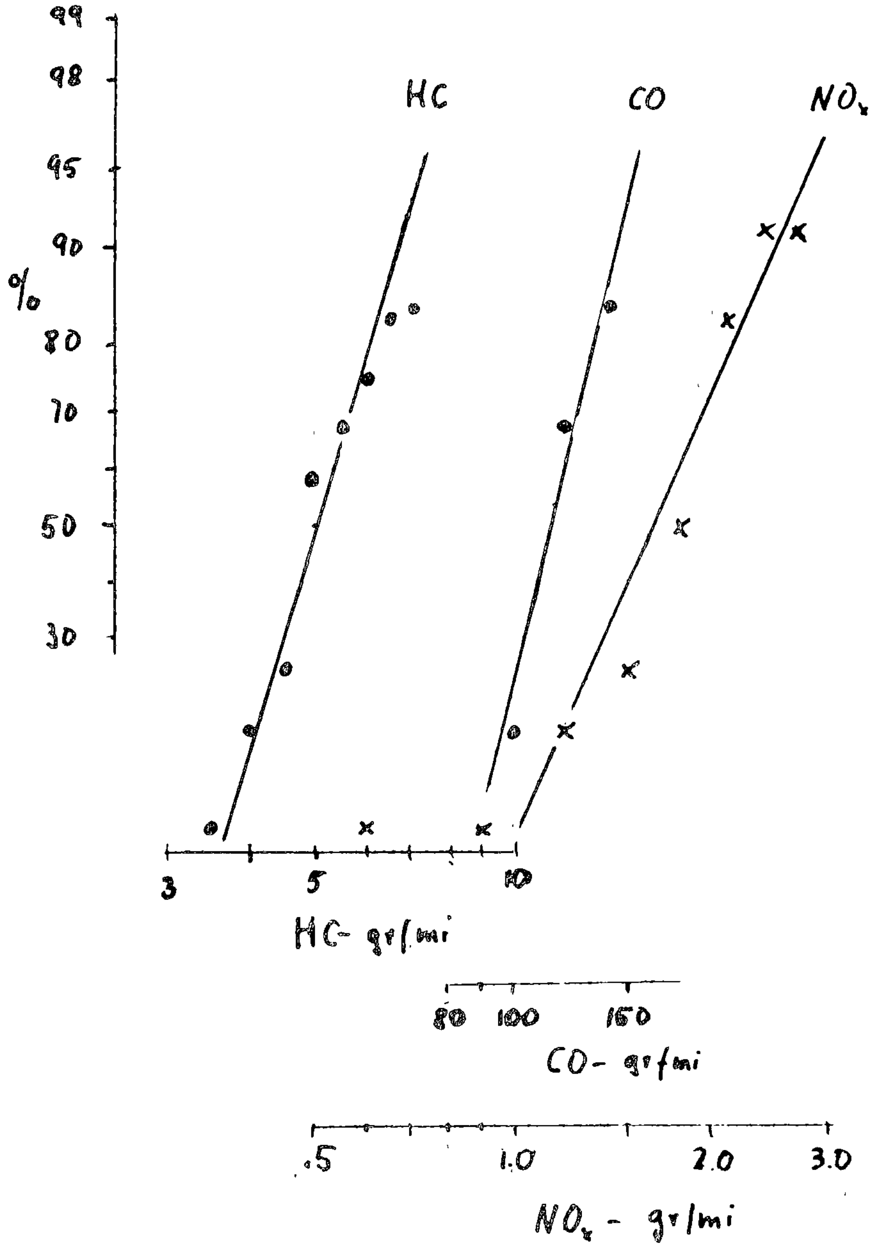


Figure 8 - Cumulative Relative Frequency Distributions of Emission Constituents of Two New M-151 Jeeps, Production 1971, Tested by EPA (CVS-C Cycle) - Reference 3

Exhaust Constituent	Manufacturer Recommended Mean and Standard Deviation		1976 Standard	Computed from $\bar{x}$ and s		
	Mean $\bar{x}$	Standard Deviation s	$x_{99.9}$	Median x_{50}	x_{84}	x_{16}
HC	.183	.07	.41	.171	.247	.118
CO	.99	.52	3.4	.876	1.435	.535
NO_x	.27	.0523	.40	.265	.321	.219

Table 2

Data x_{50}, x_{84}, and x_{16} of Log-Distribution, Computed from Emission Data $\bar{x}$ and s Recommended by Unidentified Manufacturer

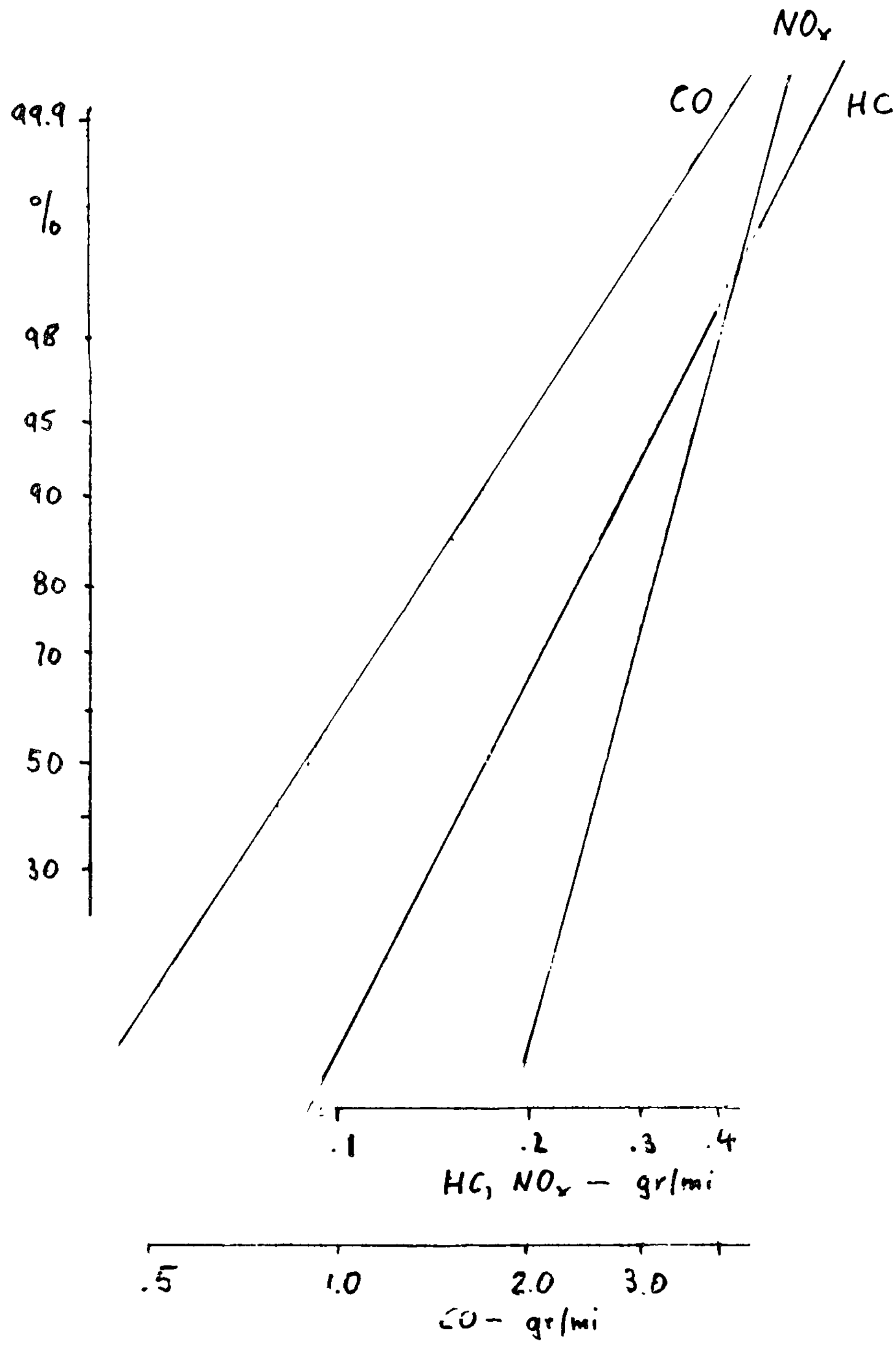


Figure 9 - Cumulative Relative Frequency Distribution of Emission Constituents, Derived from a Forecast of Unidentified Manufacturer for Model 1976 Cars

Table 3

Emission Results - Rebels - EPA
(Reference 5)

Car Number	1972 FTP grams/mile		
	HC	CO	NO_x
44044	4.23	55.38	8.04
54378	2.67	20.70	7.27
54368	2.24	13.78	6.04
54369	3.62	28.98	7.59
54380	1.60	19.52	3.59
54373	2.34	9.09	8.28
54360	1.89	9.95	6.19
54364	2.18	18.48	6.85
54370	3.28	23.31	7.82

Table 4

Emission Results - Falcons - EPA
(Reference 5)

Car Number	1972 FTP grams/mile		
	HC	CO	NO_x
54356	3.17	11.01	9.73
46982	4.19	13.07	9.96
49738	3.66	13.05	9.20
54352	2.81	16.82	8.02
49741	4.21	24.62	12.17
46983	3.33	22.97	8.06
46988	4.34	13.67	10.49
46985	3.56	15.77	9.31
49734	4.04	12.91	7.85

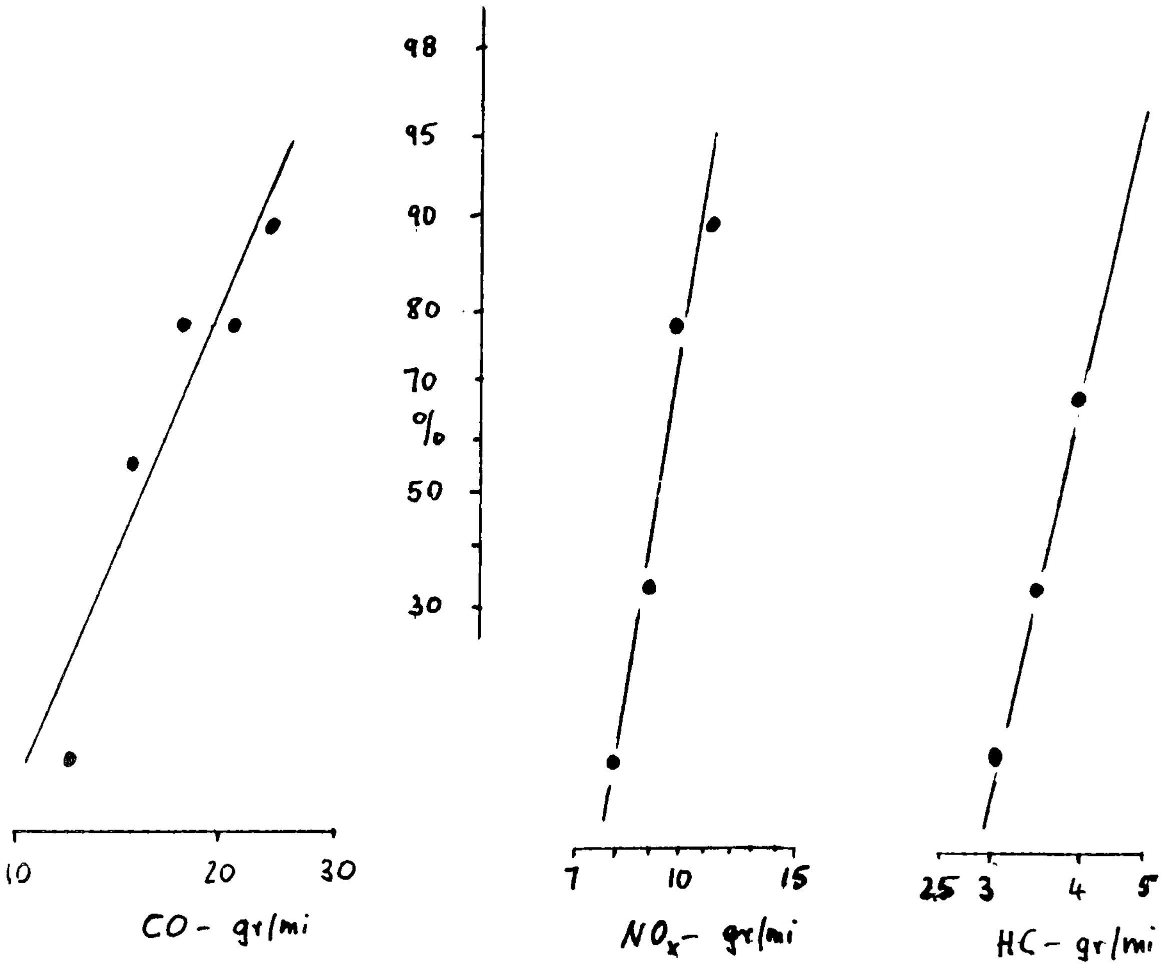


Figure 10 - Cumulative Relative Frequency Distribution of Emission Constituents of Nine Falcons, Model 1970, Tested by EPA (CVS-C Cycle) - Reference 5

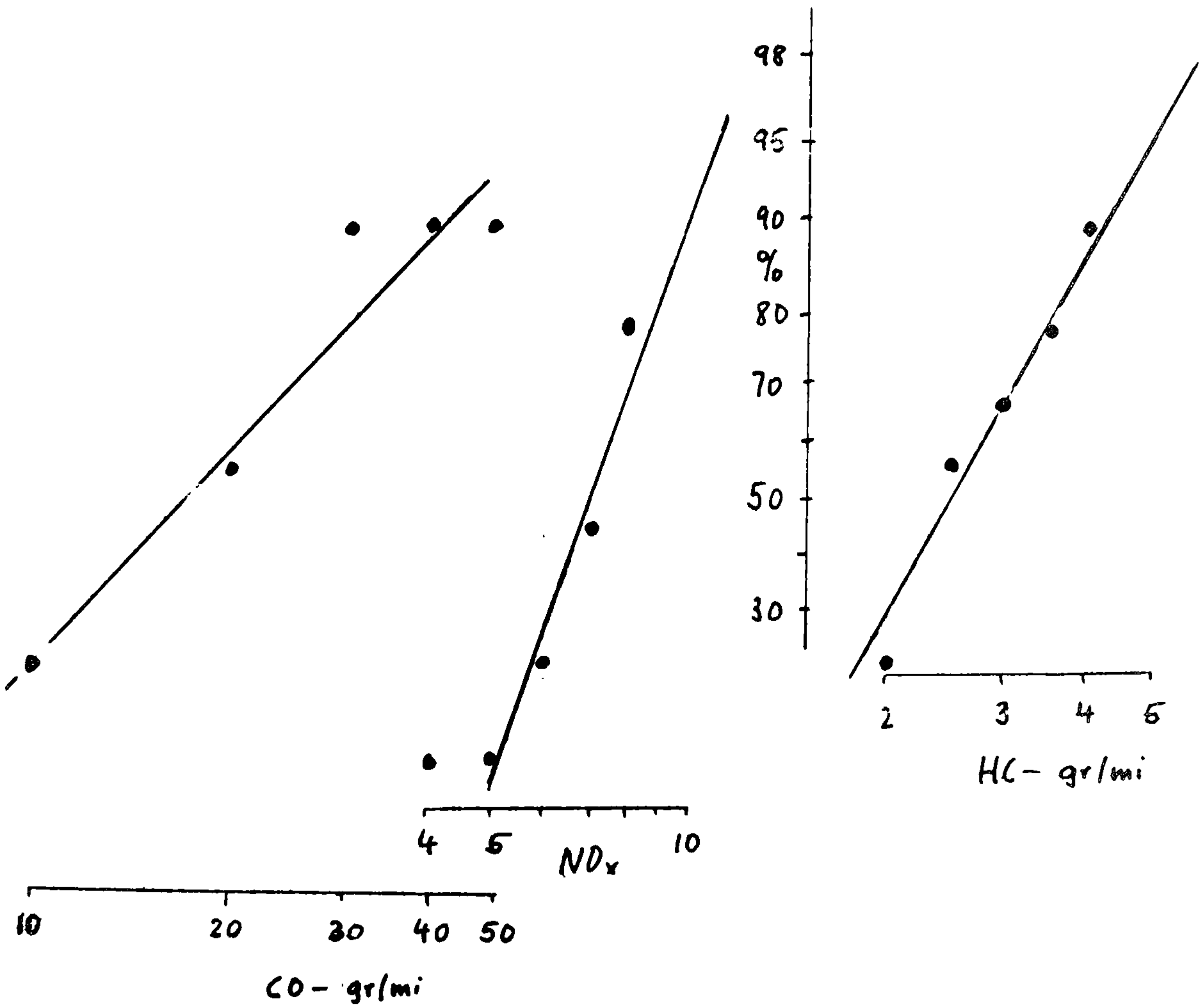


Figure 11 - Cumulative Relative Frequency Distribution of Emission Constituents of Nine Rebels, Model 1970, Tested by EPA (CVS-C Cycle) - Reference 5

Table 5

Emission Data of Two Unidentified Cars
Equipped with Experimental Control Systems

	HC	CO	NO_x	Mileage
Car #1	0.41	2.48	1.49	0
	0.02	2.76	2.27	1,000
	0.38	2.37	1.45	2,000
	0.35	2.36	2.17	5,000
	0.32	2.20	1.52	10,000
	0.38	3.90	2.25	15,000
Car #2	0.51	2.80	-	0
	0.21	3.20	1.90	0
	0.31	2.30	1.65	1,000
	0.04	1.50	1.90	2,000
	0.26	6.80	1.30	5,000
	0.47	2.00	1.15	5,000

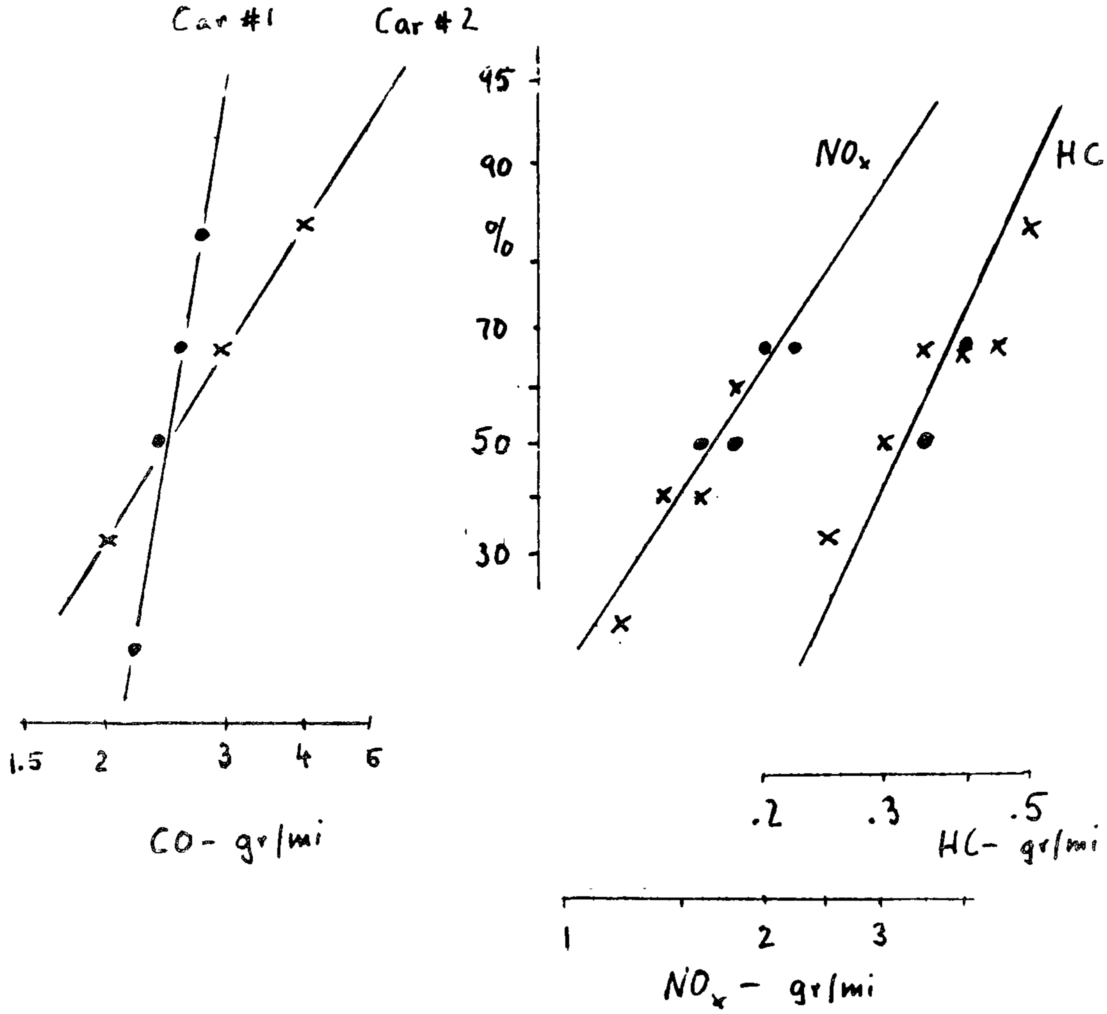


Figure 12 - Cumulative Relative Frequency Distribution of Emission Constituents of Two Unidentified Cars (CVS-C Cycle)
o = Car #1
x = Car #2

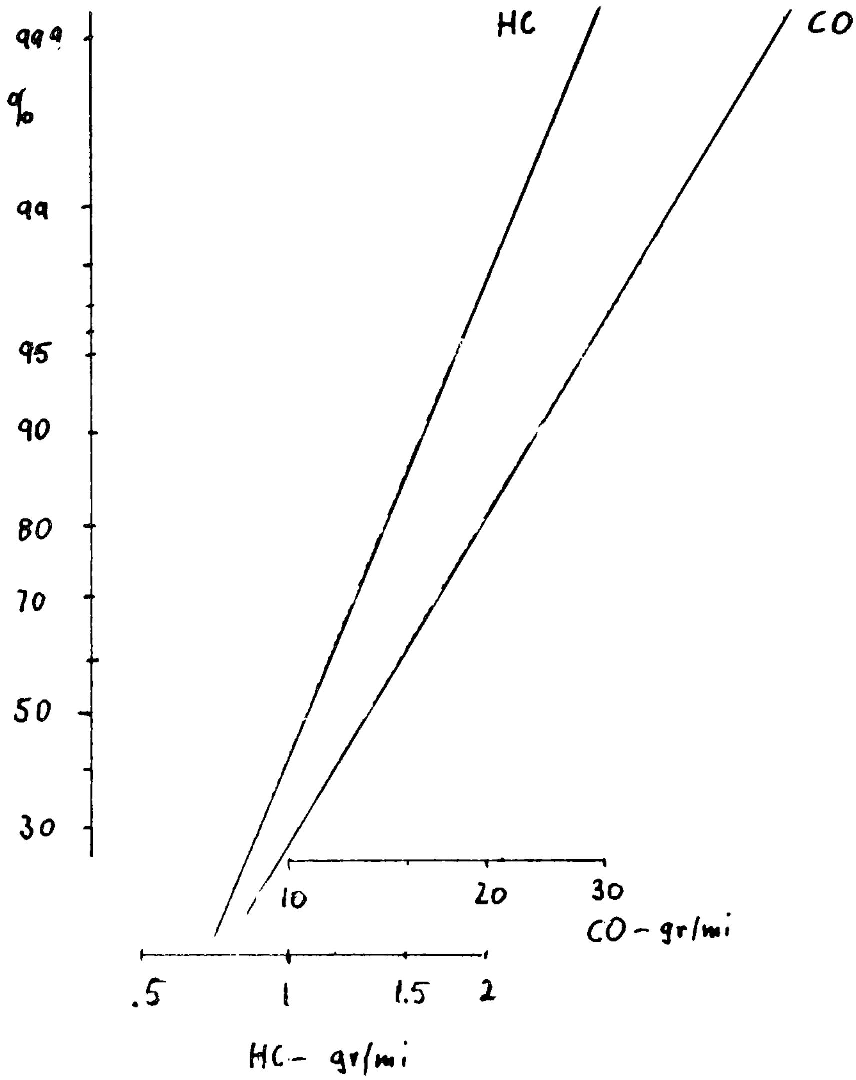


Figure 13 - Cumulative Relative Frequency Distribution of Emission Constituents of a Large Number of Unidentified Cars (FTP Hot-Cycle)

(7) In a memorandum to EPA,[7] CAL described the results of an emission test performed on a jeep equipped with a low emission FCP engine. The results of this evaluation together with the computed x_{50}, x_{16}, and x_{84} values are presented in Table 6 and plotted in Figure 14.

Table 7 summarizes the important input information of Figures 3 through 14. The first columns indicate the sources of variability involved in the data plots. The test error, discussed earlier, is always present in all measurements. The product variation enters if more than one engine is involved. Figure 11 gives an example where the emission distribution of nine Rebels is plotted and where the distribution reflects the test error of each engine, combined with the variation between the nine engines.

The deterioration factor would apply for measurements taken at different stages of engine life, and the variation between engines would add still another portion of the total variation if different engine makes are involved, such as in Figures 5 through 7.

No conclusions about these four sources of variation can be made at this point other than to state their presence or absence. We would expect the total variation to increase with the number of error sources involved, but this conjecture would have to be verified by facts. The following evaluation of test results is one first step in this direction.

Exhaust Constituent	Measured Emission Data		Computed from $\bar{x}$ and s		
	Mean $\bar{x}$	Standard Deviation s	Median x_{50}	x_{84}	x_{16}
HC	.328	.0578	.323	.385	.271
CO	.824	.1347	.813	.956	.691
NO_x	.298	.0476	.294	.345	.250

Table 6

Data x_{50}, x_{84}, and x_{16} of Log-Distribution, Computed from $\bar{x}$ and s of Measured Emission Data of FCP Engine (Reference 7)

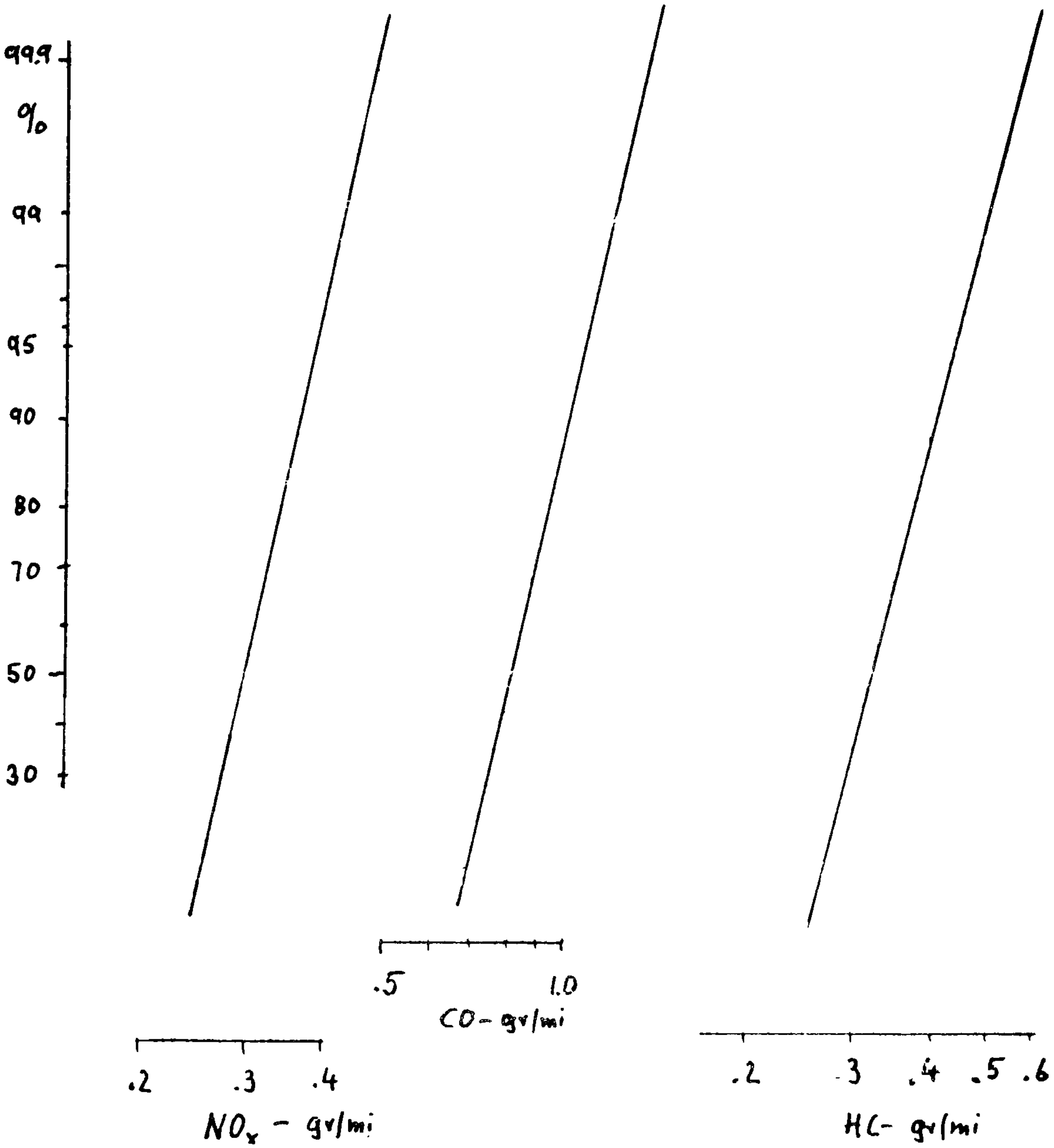


Figure 14 - Cumulative Relative Frequency Distribution of Emission Constitutents of One Jeep Equipped with Low-Emission FCP Engine (CVS-CH Cycle) - Reference 7

Figure	Test Error	Product Variation	Deteriora-tion Factor	Different Car Makes	Mileage	Emission Component HC	CO	NO_x	Test Procedure	Number of Observations	Remarks	Reference
3	x	x	x	-	0-20,000	x			FTP-hot	≈160	Unidentified GM cars (model 1970) in customer use	(1)
4							x					
5						x						
6	x	x	x	x	4,000 to 32,000		x		CVS-HC	54	Unidentified fleet cars	(2)
7								x				
8	x	-	-	-	< 60	x	x	x	CVS-C	12	Two Jeeps (production 1971) no emission controls	(3)
9	x	x	-	-	end of assem-bly line	x	x	x	CVS-HC	∞	Anticipated level of 1976 cars	(4)
10	x	x	-	-	Low	x	x	x	CVS-C	9	1970 Falcons (Ford)	(5)
11	x	x	-	-	Low	x	x	x	CVS-C	9	1970 Rebels (AM)	
12	x	-	x	-	0-12,000	x	x	x	CVS-C	6/12	Two cars equipped with exper-imental 1975 control systems	(6)
13	x	x	no info	-	no info	x	x	-	FTP-hot	"large"	1970/71 production engines of same displacement	(6)
14	x	-	-	-	4,000	x	x	x	CVS-C	≈17	FCP engine in jeep	(7)

Table 7

Summary of Input Data for Figures 3 through 14

DATA EVALUATION

In order to obtain some idea of the relative magnitude of the test error and its relation to emission differences due to product variation and other sources, all distributions pictured in Figures 3 through 14 were standardized. For normal distributions, standardization can be achieved by referring the standard deviation, s, to the mean, $\bar{x}$. The relative variation, $s/\bar{x}$, would then serve to compare the various normal distributions on a percentage basis. For logarithmic distributions, as they prevail in emission tests, comparisons can be made if all cumulative frequencies are referred to the median value, x_{50}. This is accomplished easily by shifting the averaging distribution lines transversely to the emission value "unity", as indicated in Figure 15.

In this fashion, all emission distributions, Figures 3 through 14, were standardized. Figure 16 shows the standardized HC distributions, Figure 17 those of CO, Figure 18 of NOx. Each distribution is represented by a straight line. Note, however, that these lines reflect distributions of samples only and not of populations. Table 7 indicates that, in some instances, the number of observations per sample was rather small, so that the deviation of the sample distribution from the population distribution may be large. Because the distributions are not normal, we cannot readily compute their confidence limits. We have to keep in mind, however, that the positions of distribution lines for small samples (Nos. 8, 10, 11, 12, and 14) are not exact; they can deviate towards smaller values by, say, 20% or more, and towards larger values by, say, 50% or more.

In general, one would surmise the slope of the cumulative emission distribution to reflect the number of involved error sources; emission data

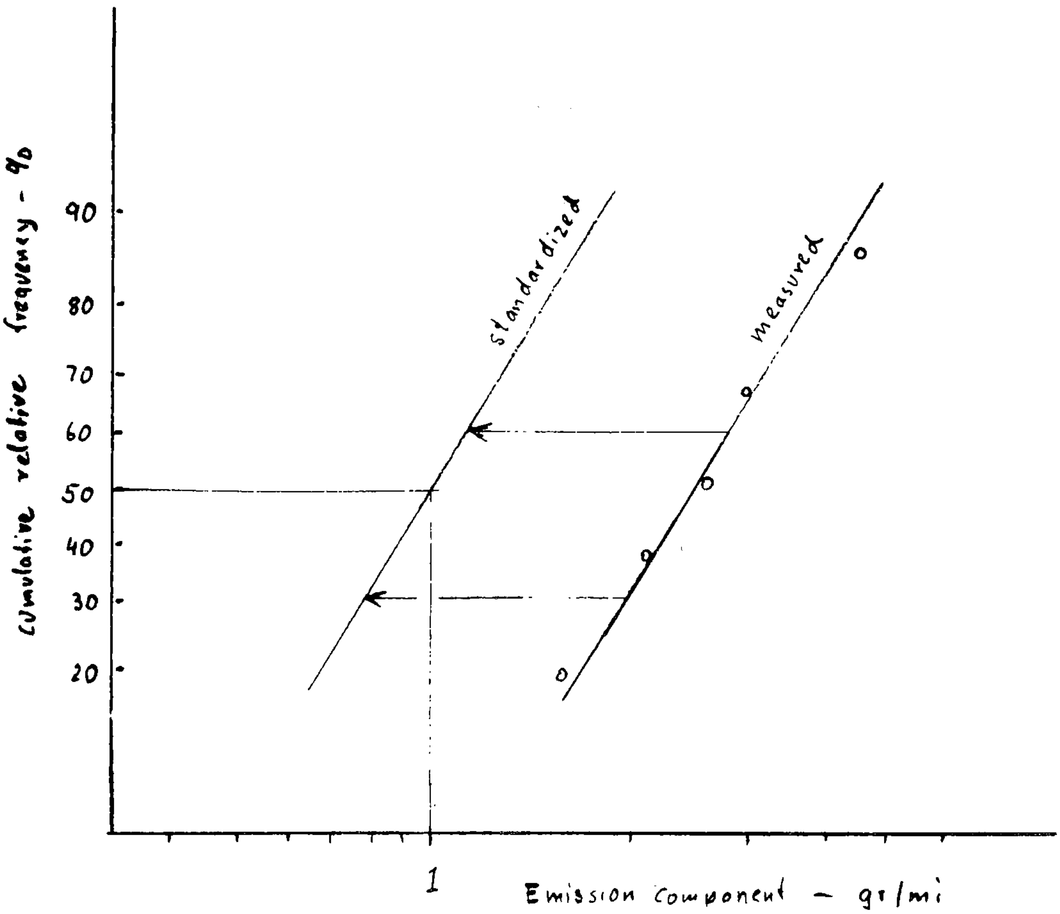


Figure 15 - Standardization of Cumulative Relative Frequency Distribution

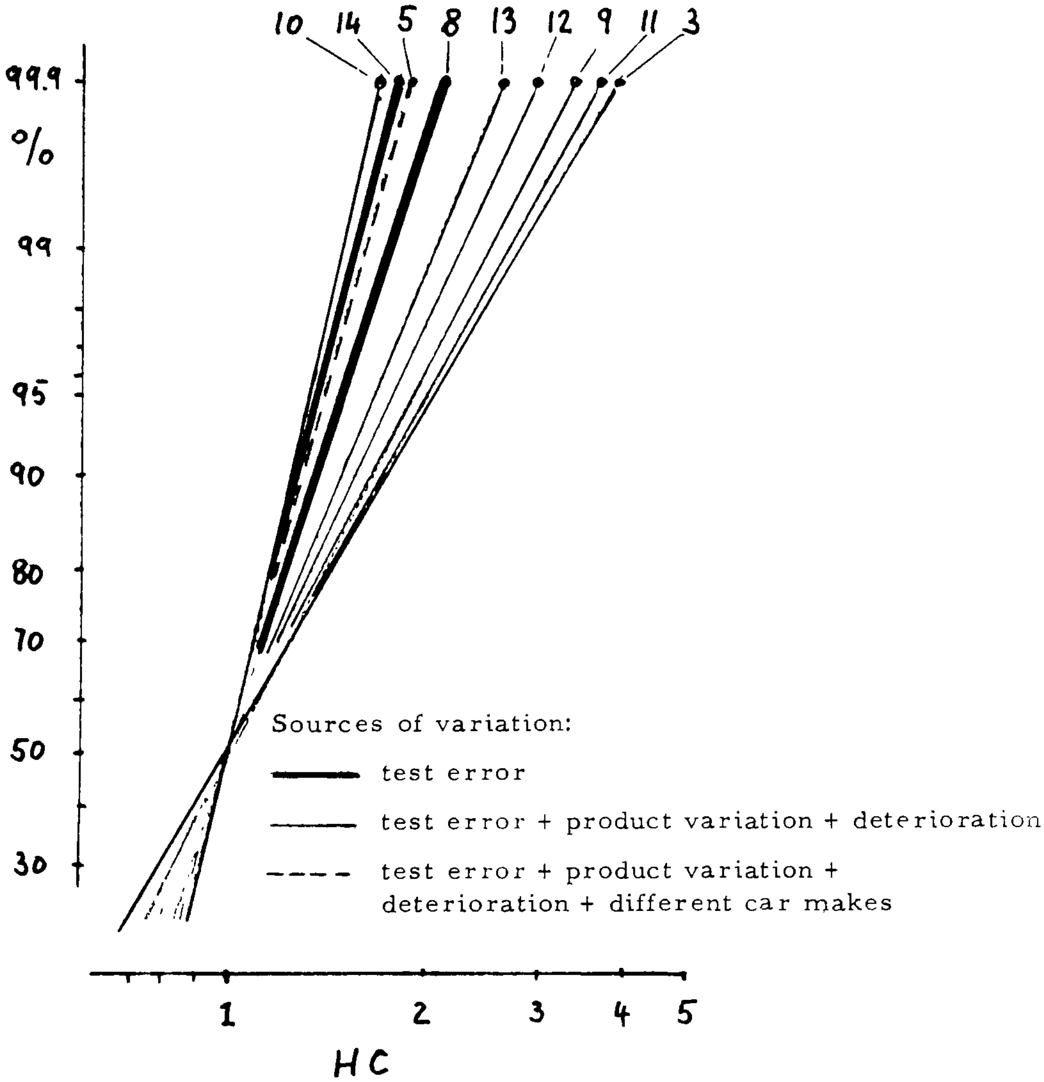


Figure 16 - Cumulative Relative Frequency Distributions of HC Emissions -- Standardized. (The numbers refer to Figures 3 through 14)

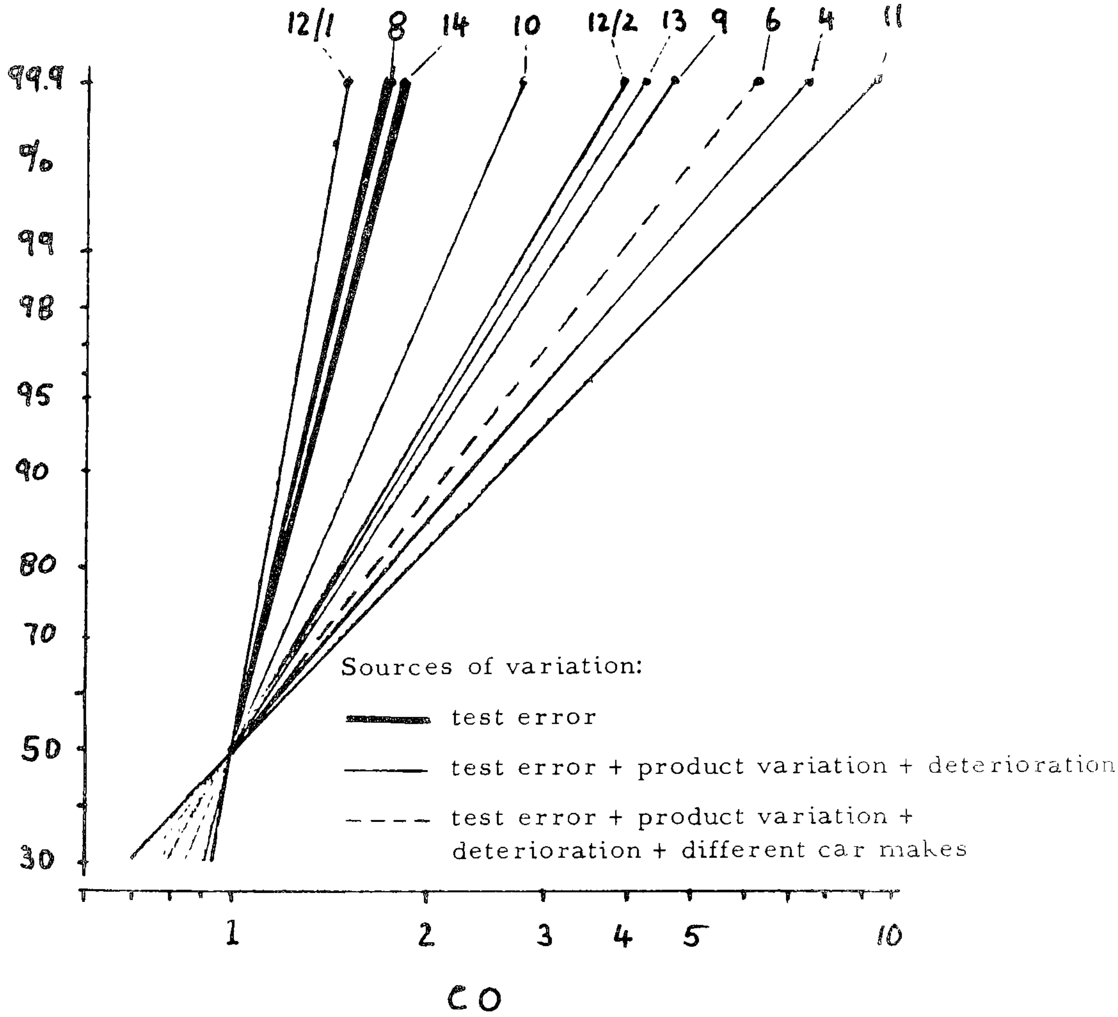


Figure 17 - Cumulative Relative Frequency Distributions of CO Emissions -- Standardized. (The numbers refer to Figures 3 through 14)

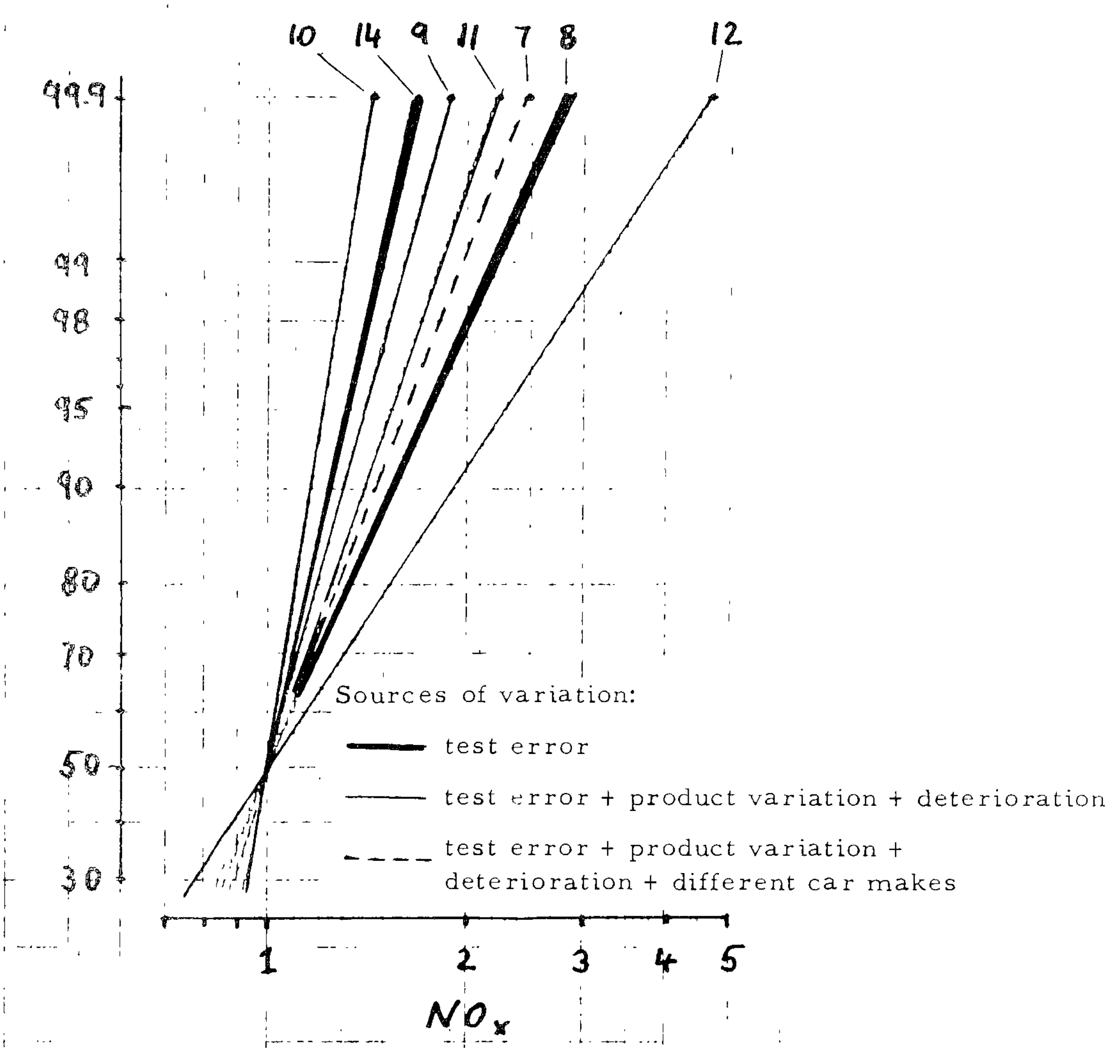


Figure 18 Cumulative Relative Frequency Distributions of No_x Emissions -- Standardized. (The numbers refer to Figures 3 through 14)

encompassing the test error, the product variations, the deterioration due to mileage, and the variation between cars should be larger than those containing only, say. the test error. Figures 16 through 18, however, do not always comply with this idealized picture. The lowest CO variation is realized by a car equipped with 1975 experimental control equipment, curve No. 12/1 in Figure 17, but the same car exhibits the highest NOx variation, curve No. 12 in Figure 18; whereas the lowest NOx (and HC) variation is achieved by a group of nine new Falcons, model 1970, curve No. 10 in Figures 16 and 18, and not by a single engine.

The largest HC variation is exhibited by approximately 160 GM cars, model 1970, driven by various customers up to 20,000 miles, curve No. 3 in Figure 16. This is as expected because the emission data of the GM cars encompass all error sources that can possibly be involved. The largest CO variation, on the other hand, is realized by a group of nine new Rebels, model 1970, curve No. 11 in Figure 17, with only the test error and the variability between cars involved.

Figures 16 through 18 display also the distributions of two single engines. Curve No. 8 shows the distribution of a standard L-141 engine in a M-151 military jeep, production 1971, with practically no emission controls, and curve No. 14 the distribution of a FCP 141 engine in the same jeep with maximum emission controls. The CO and the HC variation of both curves are low (albeit not the lowest) and essentially equal, Figures 16 and 17. The NOx variation of the FCP engine, curve No. 14 in Figure 18, is still low, but that of the uncontrolled engine, curve No. 8, is high.

We suspect this lack of consistency to be caused primarily by the test error. It seems likely that the test error, which is present in all measurements

changes widely with the car make, the car service, the measuring equipment, the test operator, and the like, to the extent as to obscure the contributions of other sources of variation. The Automobile Manufacturers Association recently conducted a survey[8] of six emission test laboratories by having each of them test the same low emission vehicle. The results are listed in Table 8. Although the given data do not allow computation of the test error within labs, we suspect the variation between labs to be large.
For instance, the test error of the Ford lab seems to be much larger than that of the AMC lab as regards HC emissions.

Similar differences of test errors may occur when cars of different makes are tested. The Falcons, model 1970, for instance, exhibit a relatively small variation in HC and NOx, see distribution No. 10 in Figures 16 and 18. Since this variation encompasses both the test error and the variation between cars, the reproduction error must be even smaller than the shown total variation. The experimental cars, distribution No. 12 in Figures 16 and 18, on the other hand, show a relatively large test error in HC and NOx. The ratio between the reproduction errors of the Falcons and the experimental cars is at least 1.8 for HC and 3.2 for NOx. It is impossible to decide to what extent these large ranges are caused by the cars themselves and not by errors of the instrumentation, but we suspect that a large part is due to differences in car performance.

All we can state then is that the test error is by no means a constant. It depends on the measuring equipment, on the test procedure, on the car make, on the operator, and the like. The test error can be small as evidenced by distribution No. 10 in Figures 16 and 18, and No. 12/1 in Figure 17.
If we want to assign numbers, we could form the ratio of $x_{99.9}$ and x_{50} (the median) and state that since the ratio of $x_{99.9}/x_{50}$, encompassing test

Test Laboratory	No. of Tests	1972 Test Procedure Grams/Mile (Average)		
		HC	CO	NO_x
American Motors Corporation	2	0.48	22.93	1.99
General Motors Proving Ground	1	0.58	23.00	2.03
Chrysler	1	0.43	33.26	1.11
Ford	3	1.00	22.84	1.54
International Harvester	2	0.82	30.60	3.08
Environmental Protection Agency	1	0.80	24.06	1.67

Table 8

Emission Data of the Same Vehicle Tested at Various Laboratories

error plus variations between cars, reached a minimum of approximately 1.5, the test error could well be significantly smaller. In other words, the smallest test error encountered in our studies was of a magnitude that would refer 99.9% or more of all data (measured repeatedly on one single engine under identically controlled conditions) below the limit of 1.5 x median value. Whether this small test error can be realized for all emission measurements if performed with care and good instruments, or whether some engines perform erratically to the extent as to surpass this minimum even with ideal instrumentation, cannot be determined without additional information not available at this time.

Figures 16 through 18 exhibit also the maximum variation of emission data. Maximum variation should be expected at a combination of all possible sources of variation, i.e., the test error, product variation, deterioration factor, and variation between cars of different makes, which is realized by distribution No. 5 in Figure 16, No. 6 in Figure 17, and No. 7 in Figure 18. The data show, however, that distribution No. 5, which reflects the distribution of HC data of 54 fleet cars of different makes, is rather small compared to distribution No. 3 reflecting the HC distribution of 160 GM cars. This is perhaps understandable because fleet cars are usually better serviced than cars in customer use. The CO data of both groups of car, however, are almost equal, see distributions No. 6 and 4 in Figure 17. So again, no consistent picture evolves.

On the whole, it seems that the maximum variation of CO data is about twice as large as the maximum variation of HC and NOx data. Expressed in terms of the ratio of $x_{99.9}/x_{50}$, the maximum range is

$$x_{99.9}/x_{50} \gtrsim \begin{cases} 4 & \text{for HC} \\ 9 & \text{for CO} \\ 4.5 & \text{for } NO_x \end{cases}$$

As regards the FCP engine, we may conclude that although its test error is low, see distribution No. 14 in Figures 15 through 17, it is not the lowest. There seems no reason to assume that the FCP engine would exhibit a lower test error than any ordinary carbureted engine.

SUMMARY

Emission data of various groups of vehicles, all equipped with internal combustion engines, were analyzed with the objective of extracting information about the test error of the single car tested repeatedly with the same equipment under identically controlled conditions. The analysis revealed that:

- The distribution of emission data can be transformed to approximate normality by using the logarithm of the emission values rather than the values themselves.
- It seems that the test error changes widely between cars and between test labs.
- The smallest test error encountered was of a magnitude that would place 99.9% of all measured HC, CO, and NOx data below the limit of 1.5 x median value. Whether this range is typical for good measuring techniques and well-serviced engines could not be decided.
- There seems no reason for believing that the test error of the FCP engine would yield a lower test error than ordinary carbureted engines.

REFERENCES

(1) General Motors Corporation: "A Progress Report by the General Motors Corporation to U.S. Environmental Protection Agency" - March 12, 1971.

(2) Private communication

(3) Environmental Protection Agency, Office of Air Programs: "A Report on the Exhaust Emissions of the 1971 Production Version of the Army M-151-Jeep" - by J.C. Thomson, Rep. No. 71-27, April 1971.

(4) Private communication

(5) HEW, Nat. Air Pollution Control Administration, Div. of Motor Vehicle R & D: "Exhaust Emissions from Ten GSA Rebels and Ten GSA Falcons Equipped with LPG Conversion Kits" by H. L. Gompf, Rep. No. 71-10, October 1970.

(6) Private communication

(7) Cornell Aeronautical Laboratory: "Analysis of Emissions Test Data on Low-Emission FCP Engine Installed in M-151 Vehicle" - prepared for EPA, December 1971.

(8) Ford Motor Company: "1975/1976 Light Duty Vehicle Emission Control Program" - Status Report to the Environmental Protection Agency, October 18, 1971.

APPENDIX E EMISSIONS ANALYSIS OF 54 FLEET-TYPE PASSENGER VEHICLES

D. J. Schuring

INTRODUCTION

Fifty-four 1970-model fleet cars of different makes and mileages were tested by EPA using both the CVS cold-start and the CVS hot-start test conditions. On November 15, 1971, CAL received from EPA the results of these emission tests with the following remarks: "A two-bag cold start test was run on each car followed by a two-bag hot start test. To obtain three-bag data, it is necessary to weight and assemble the data from the two cold start bags with the first bag of the hot start test using the procedure described in the Federal Register." The identity of the vehicles could not be disclosed by EPA in any way other than by range of engine displacement, number of cylinders, accumulated mileage, and car weight.

The emission data listed in the computer print-outs gave occasion to study the following questions:

- To what extent does the accumulated mileage influence the emissions level?
- Is there a significant difference between the emissions levels for cold start and hot start?
- What is the numerical relation between the CVS-C and the CVS-CH testing procedure?

The answers to these questions were felt to provide, by inference, some inputs into the planning of Stratified-Charge-Engine tests.

INFLUENCE OF ACCUMULATED MILEAGE ON EMISSIONS LEVEL

The accumulated mileage of the cars tested ranged from approximately 4, 000 miles to 35, 000 miles. A visual inspection of the emissions data from an analysis of each bag plotted versus accumulated mileage strongly suggested negligible correlation between the two variables, Figures 1 through 12. An analysis of variance, described elsewhere, also failed to reveal a mileage trend (nor did it confirm any other trend associated with number of cylinders, displacement, and car weight). Hence, all cars can be considered as drawn from the same homogeneous population.

As an interesting aside, it may be mentioned that in contrast to the toxic emission components (CO, HC, and NO_x), the non-toxic component, CO_2, did indeed reveal a mileage correlation. Figure 13 shows CO_2 data in gr/mi for the first cold bag as a function of mileage suggesting an upward trend of emissions with mileage. An analysis of variance, Table 1, confirmed the significant contribution of mileage to the overall variability of CO_2.

Table 1

Analysis of Variance on CO_2 Emissions (gr/mi) of Fifty-four Cars, Divided into Three Mileage Classes ($\leq$ 10, 000; 10, 000 - 20, 000; > 20, 000 miles)

Source	Sum of Squares	Degree of Freedom	Mean Square
Between mileage classes	14.92×10^4	2	7.69×10^4
Within mileage classes	65.80×10^4	51	1.29×10^4
TOTAL	81.72×10^4	53	$F = 6.17$ $F_{2, 51, 99\%} \approx 5.0$

INFLUENCE OF START CONDITIONS ON EMISSIONS LEVEL

Figures 14 through 17 depict the means and the standard deviations of the HC, CO, NO_x, and CO_2 constituents for each of the four bags. The mean is indicated by a small circle, the standard deviation by the length of the vertical bar extending from the mean value in both directions. Obviously, the HC and the CO contents (in gr/mi) of the first (cold-transient) bag are significantly higher than those of the rest. In fact, the differences between the second (cold-stabilized), third (hot-transient), and fourth (hot-stabilized) bag are so slight as to suggest a constant level of CO and HC emissions after the first bag. Thus, the HC level of the first bag is approximately 1.6 times higher than the average level of subsequent bags, and the CO level of the first bag approximately 3.4 times higher. The situation is different for NO_x. Here, the NO_x content of the first bag is always higher than that of the subsequent bag regardless of whether the engine is started after a cold soak or a hot soak. The CO_2 level is of lesser importance. We note that it remains approximately unchanged.

Another descriptor of interest is the standard deviation. In HC and CO emissions, the standard deviation seems to change proportionally with the mean, an observation confirmed by analyses of aircraft emission data (Reference 1). The standard deviations of both NO_x and CO_2, however, indicate no dependence on the mean.

For comparison, the emission data of the 54 cars were contrasted with the emission data of an FCP engine installed in a jeep (Reference 2). The FCP data are listed in Table 2 and indicated in Figures 14 through 17 by black bars; the bar center is identical with the mean, the bar length with twice the standard deviation. The general trends of both plots are similar although the HC, CO, and NO_x values of the FCP engine are strikingly lower than those of the 54 cars (due to effective emission control and lack of variance between engines).

Table 2

Constant Volume Sampler Results of FCP Engine
EPA, October 7, 1971 (Reference 2)

Bag	Size	Gas	Mean gr/mi	Standard Deviation gr/mi	Rel. Variation %
Cold Start	14	HC - FID	0.97	0.16	16
		CO - IR	1.30	0.43	33
		CO_2 - IR	480.	45.	9
		NO_x - CI	0.40	0.13	33
Stabilized	14	HC - FID	0.18	0.05	3
		CO - IR	0.78	0.28	36
		CO_2 - IR	509.	37.	7
		NO_x - CI	0.28	0.06	22
Hot Start	14	HC - FID	0.27	0.1	11
		CO - IR	0.76	0.26	34
		CO_2 - IR	451.	53.	12
		NO_x - CI	0.39	0.12	30

CVS-C VERSUS CVS-CH TEST PROCEDURE

The Constant Volume Sampling Procedure for 1972 prescribes a 1369 second, 7.5 mile, non-repetitive driving cycle with a 12-hour cold soak before testing and a cold start (Notation: CVS-C).

The Constant Volume Sampling Technique for 1975-76 prescribes the same 7.5 mile driving pattern as the 1972 procedure. The emissions of the first 505 seconds are collected in a "cold transient" bag, those of the next 864 seconds in a "stabilized" bag. After ten minutes hot soak with engine off, the driving cycle is repeated with the emissions of the first 505 seconds collected in a "hot transient" bag (Notation: CVS-CH). The first (cold transient) bag is weighted by a factor of 0.43, the second (cold stabilized) bag by a factor of 1, and the third (hot transient) bag by 0.57. All weighted emissions are then added and the sum divided by 7.5 to give the emissions in gr/mi.

In order to compare the CVS-C and the CVS-CH procedures, the four-bag data of the 54 cars (a cold-transient bag, a cold stabilized bag, a hot-transient bag, and a hot-stabilized bag) were assembled in two ways.

CVS-C $(Y_{ct} + Y_{cs})/7.5 = Y_m$ in gr/mi

CVS-CH $(0.43\,Y_{ct} + Y_{cs} + 0.57\,Y_{ht})/7.5 = Y_{wm}$ in gr/mi

Y_{ct} = mass emissions of pollutant (CO, HC, NO_x) as calculated from the cold-transient phase (bag 1) of the cold-start test, in grams

Y_{cs} = mass emissions of pollutant (CO, HC, NO_x) as calculated from the cold-stabilized phase (bag 2) of the cold-start test, in grams

Y_{ht} = mass emissions of pollutant (CO, HC, NO_x) as calculated from the hot-transient phase (bag 3) of the hot-start test, in grams

Y_m = mass emissions of pollutant in gr/mi; CVS-C test procedure

Y_{wm} = weighted mass emissions of pollutant in gr/mi; CVS-CH procedure

These computations were performed for each of the 54 cars (actually only 53 because one car's data were quite erratic and had to be omitted). For each car, the ratio of Y_{wm}/Y_m was then calculated, and the mean and the standard deviation of the distribution of these ratios computed. The results are listed in Table 3 together with the standard deviation of the mean ($s_{mean} = s/\sqrt{53}$).

Table 3
Comparison of Emission Levels Computed According to CVS-C and CVS-CH Method

Y_{wm} - CVS-CH Mass Emissions in gr/mi
Y_m = CVS-C Mass Emissions in gr/mi

Pollutant	Y_{wm}/Y_m Mean	Y_{wm}/Y_m Standard Deviation	Number of Cars	Y_{wm}/Y_m Standard Deviation of Mean
HC	0.88	0.07	53	0.01
CO	0.71	0.13	53	0.02
NO_2	1.01	0.05	53	0.01

Table 3 demonstrates clearly the attenuating effect of bag weighting as regards HC and CO. The HC level is reduced by 12% and the CO level by 29%. The NO_x level, on the other hand, remains unaffected.

SUMMARY

Fifty-four fleet cars of different makes and mileages, all 1970 models, were tested by EPA using both the 1972 CVS one-bag and the 1976 CVS three-bag test conditions. Although the accumulated mileage of the cars tested ranged from approximately 4,000 to 35,000 miles, a correlation between mileage and level of CO, HC, NO_x emissions was not detectable. (There was a slight but significant increase of CO_2 with mileage, however.) The mean and the standard deviations of the CO and HC emissions sampled in the cold-start bag were significantly higher than those of subsequent bags. The HC mean of the cold-transient bag was approximately 1.6 times higher than the average of the subsequent bags; the CO mean, 3.4 times. Also, the standard deviation of both emissions seemed to change approximately proportional with the mean. The situation was different for NO_x. Here, the mean of the transient bag was always higher than that of the stabilized bag regardless of starting conditions. The CO_2 level remained approximately constant for all bags. The results were qualitatively corroborated (except for mileage effects) by results of emission tests performed by EPA on one FCP engine installed in a jeep.

An investigation of the effect of bag weighting on the measured emission level revealed that the CVS-CH technique produces a 12% lower HC level than the CVS-C procedure and a 29% lower CO level. The NO_x level remains unaffected.

REFERENCES

1. Cornell Aeronautical Laboratory, Inc.: "Analysis of Aircraft Exhaust Emission Measurements" - CAL No. NA-5007-K, November 1971.

2. Private Communication: Computer Printout of EPA.

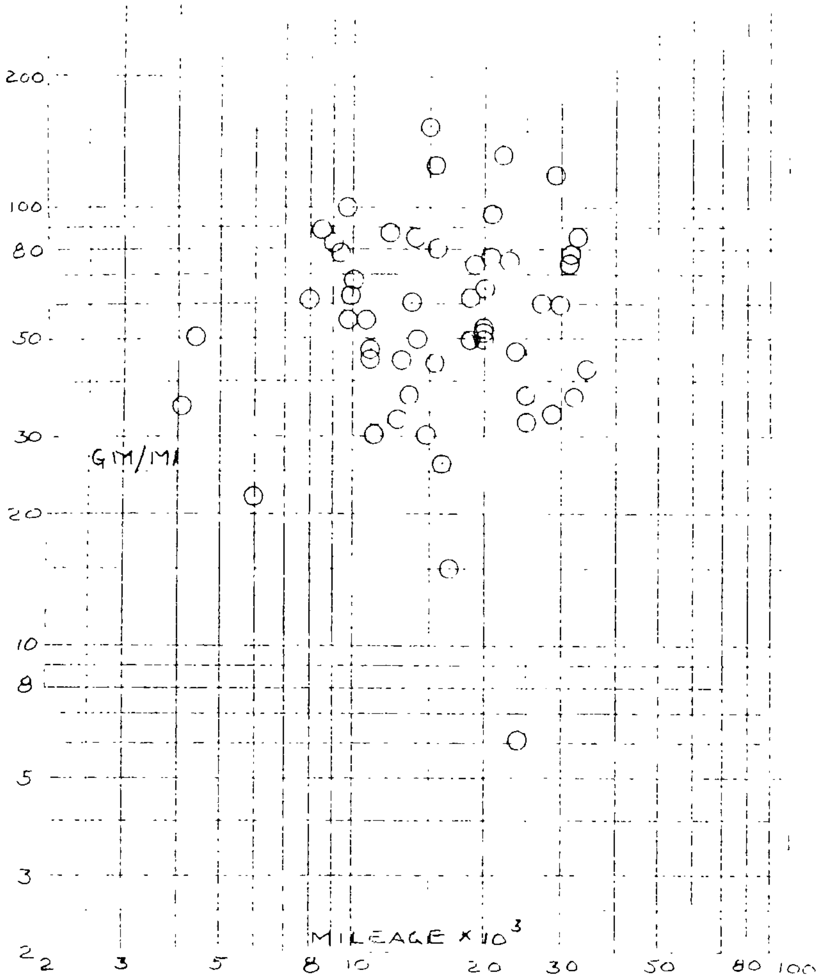


Figure 1 - CO Emissions of Cold-Transient Bag 1 Versus Accumulated Mileage

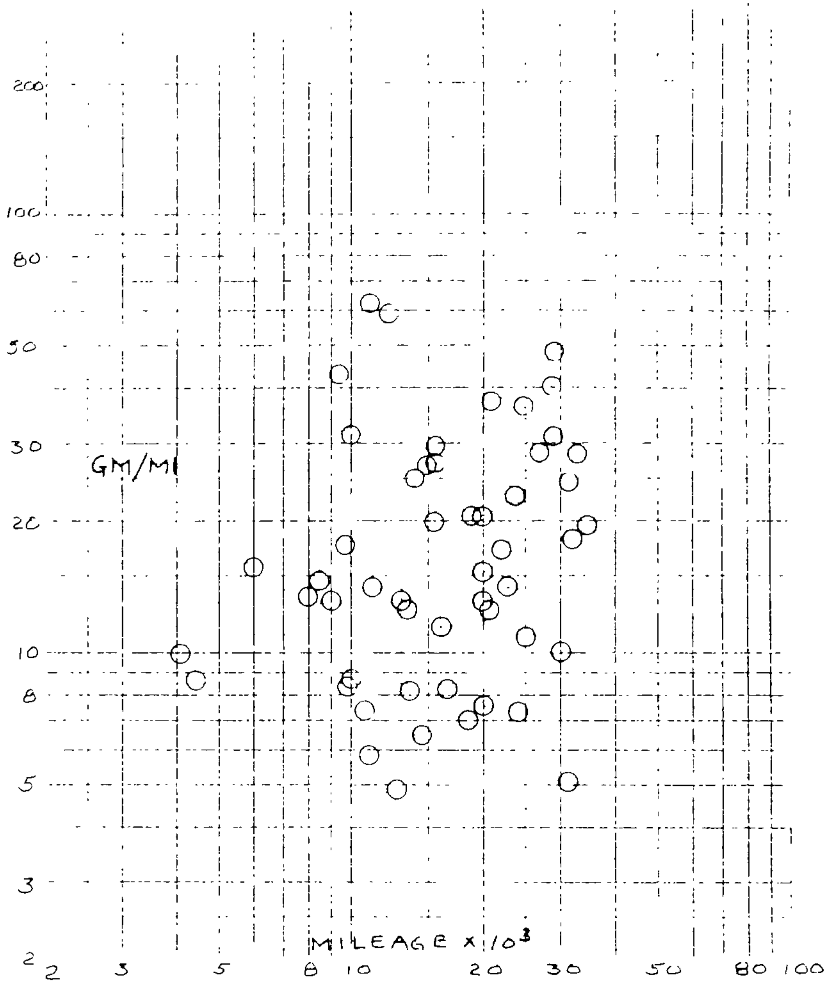


Figure 2 - CO Emissions of Cold-Stable Bag 2 Versus Accumulated Mileage

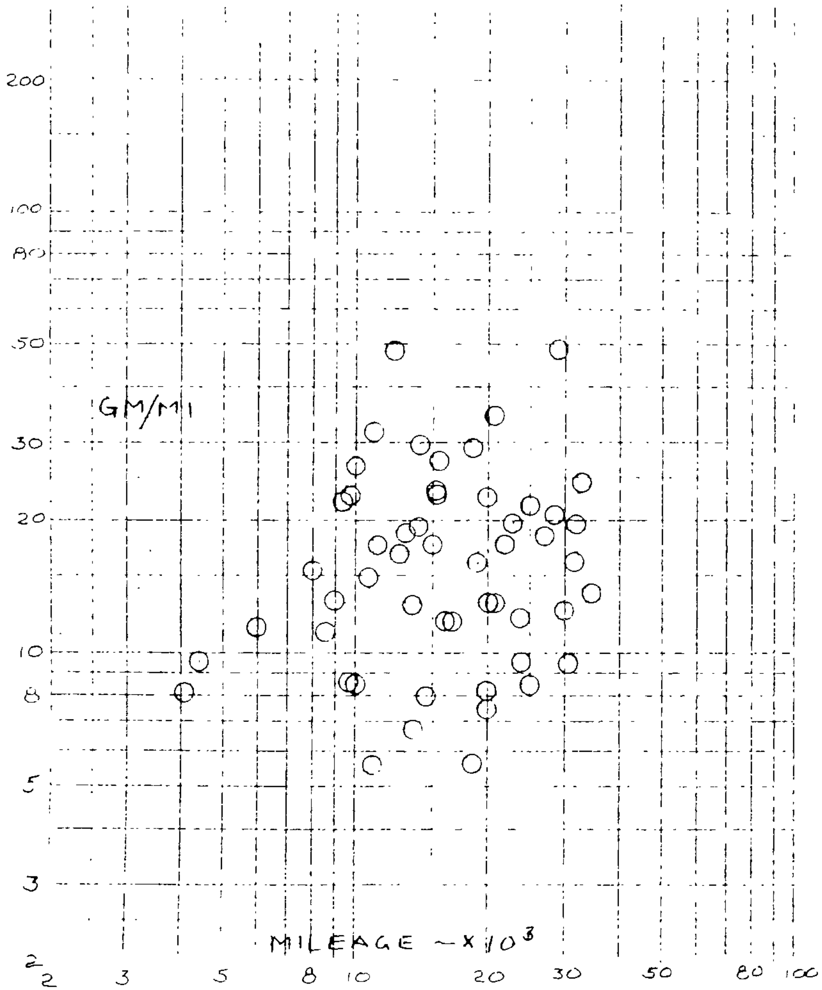


Figure 3 - CO Emissions of Hot-Transient Bag 3 Versus Accumulated Mileage

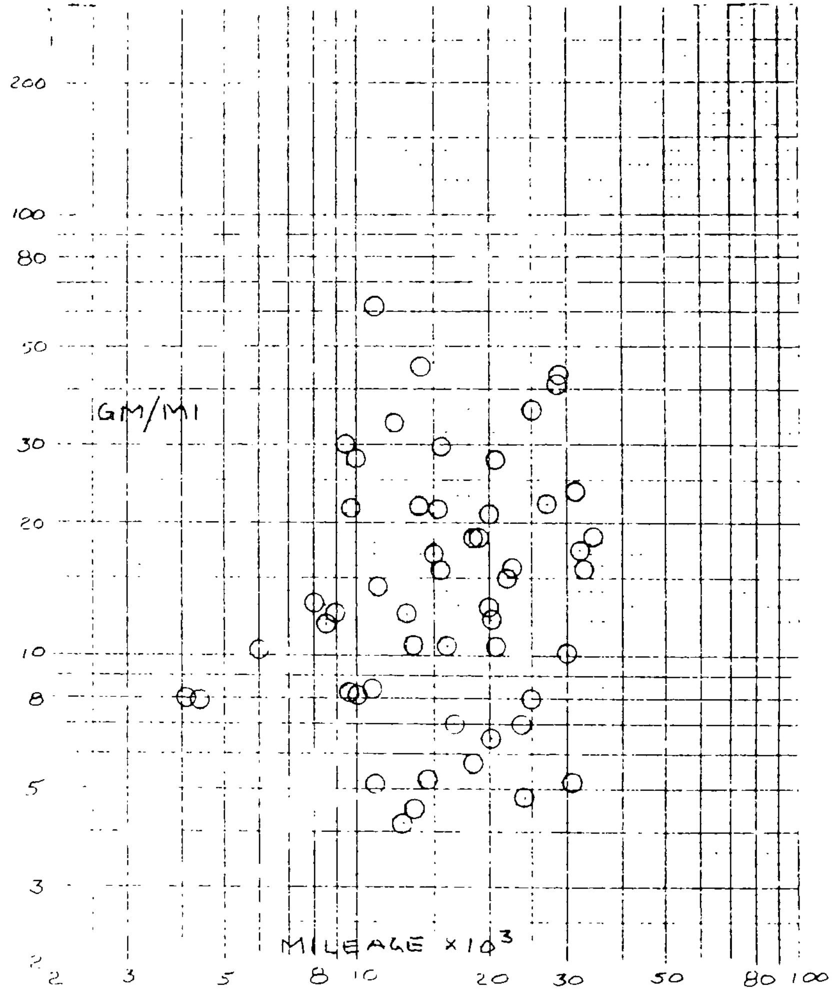


Figure 4 - CO Emissions of Hot-Stable Bag 4 Versus Accumulated Mileage

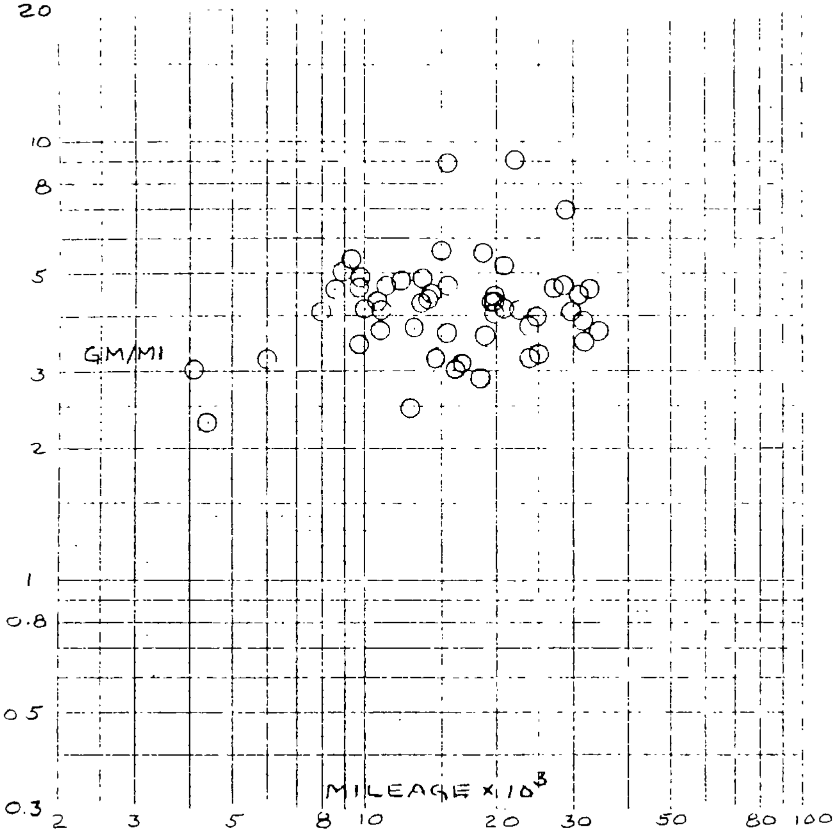


Figure 5 - HC Emissions of Cold-Transient Bag 1 Versus Accumulated Mileage

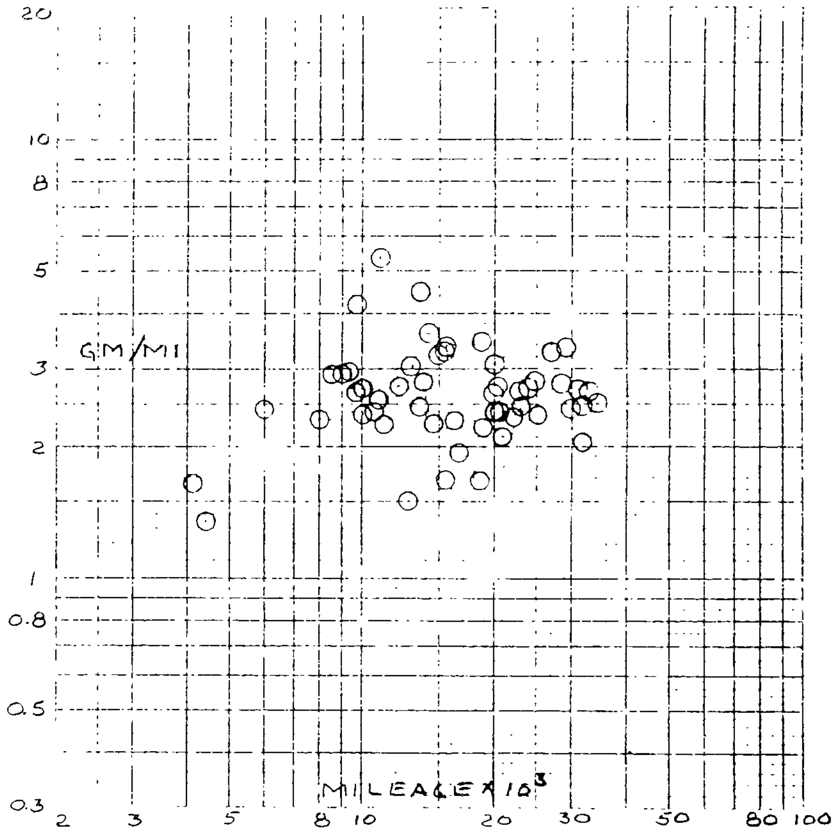


Figure 6 - HC Emissions of Cold-Stable Bag 2 Versus Accumulated Mileage

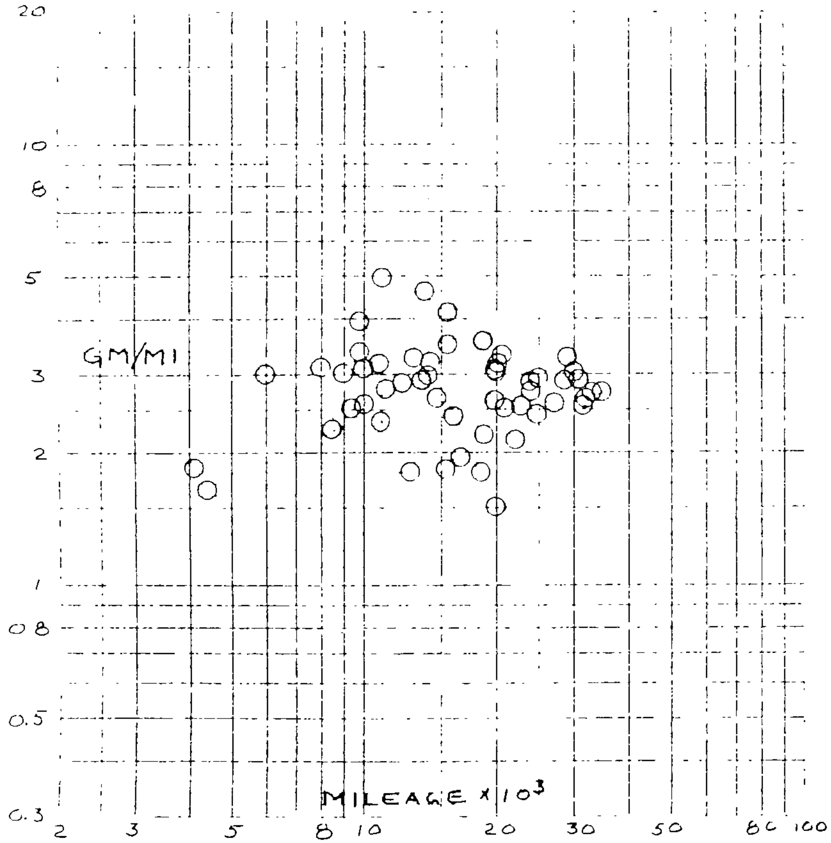


Figure 7 - HC Emissions of Hot-Transient Bag 3 Versus Accumulated Mileage

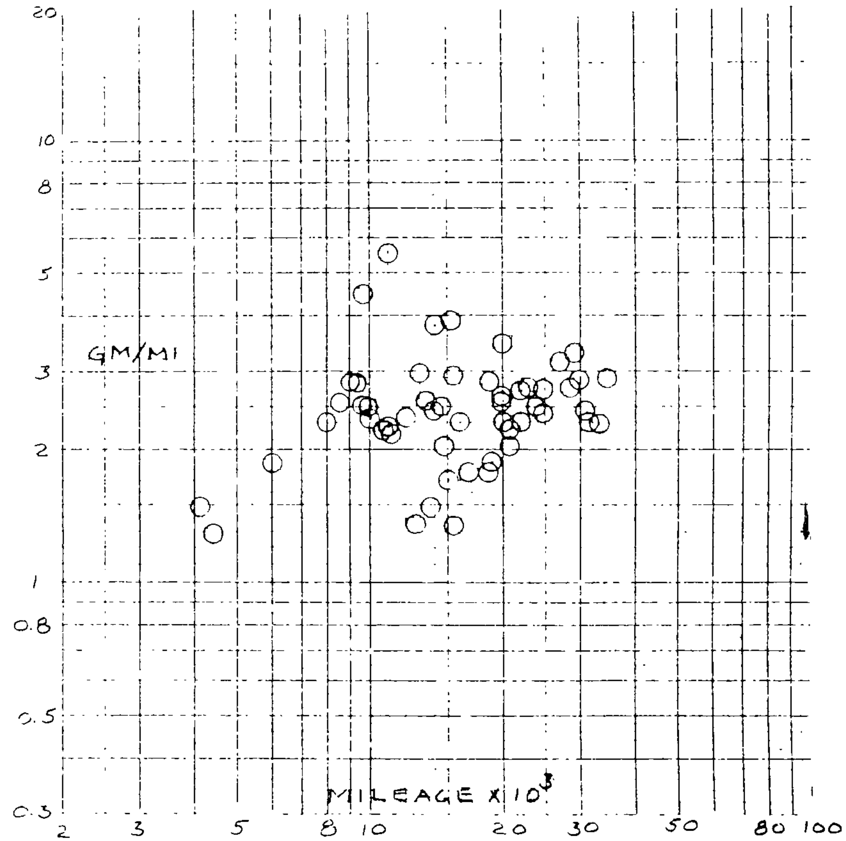


Figure 8 - HC Emissions of Hot-Stable Bag 4 Versus Accumulated Mileage

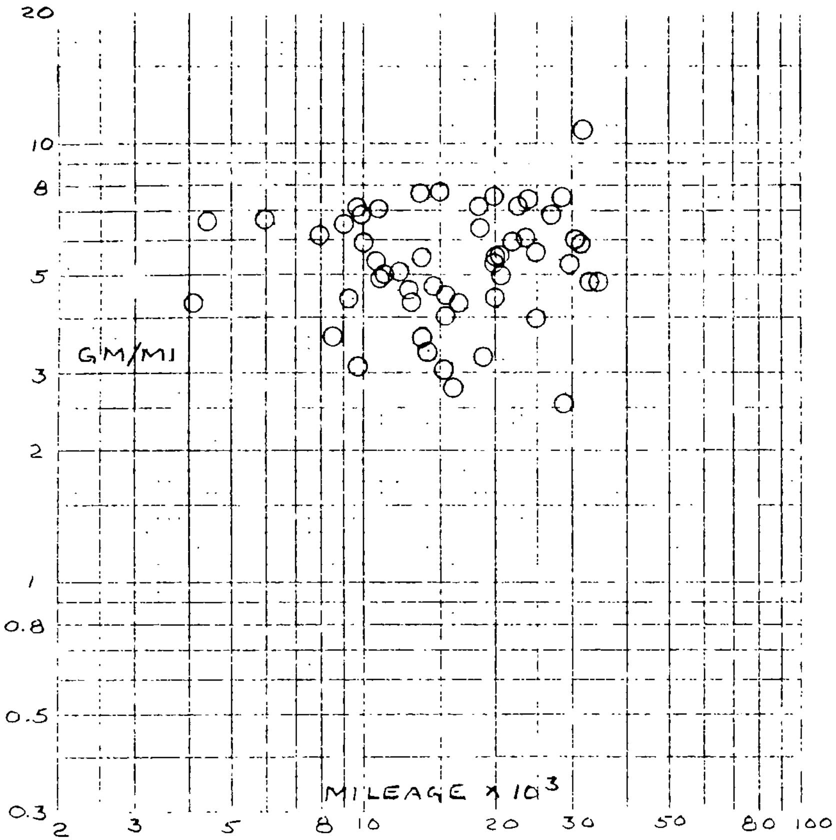


Figure 9 - NO_2 Emissions of Cold-Transient Bag 1 Versus Accumulated Mileage

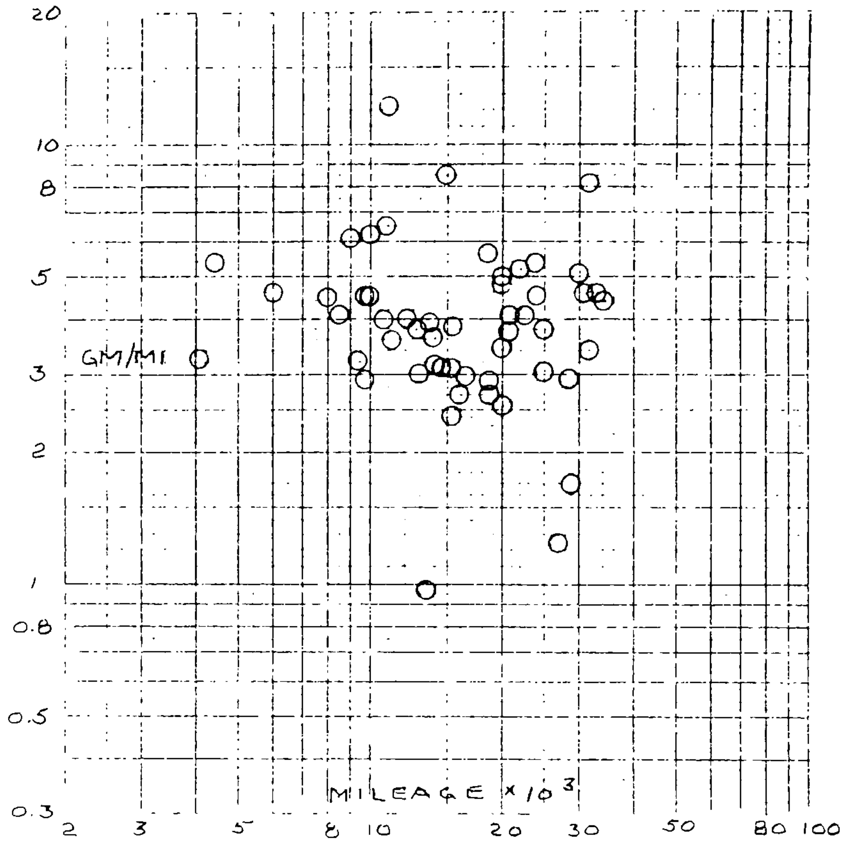


Figure 10 - NO_2 Emissions of Cold-Stable Bag 2 Versus Accumulated Mileage

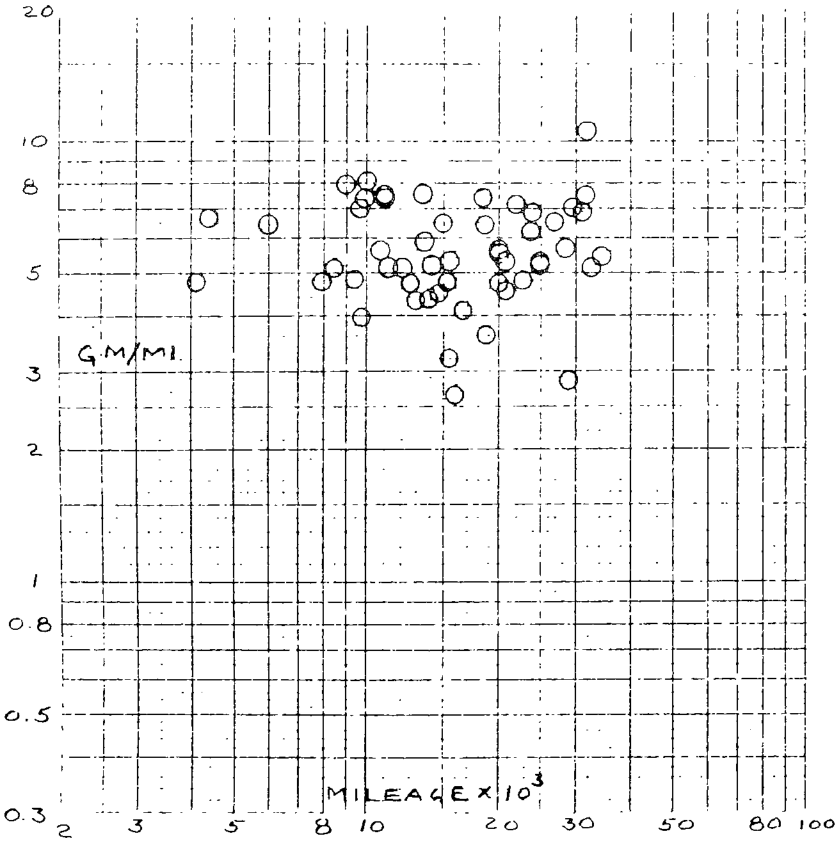


Figure 11 - NO_2 Emissions of Hot-Transient Bag 3 Versus Accumulated Mileage

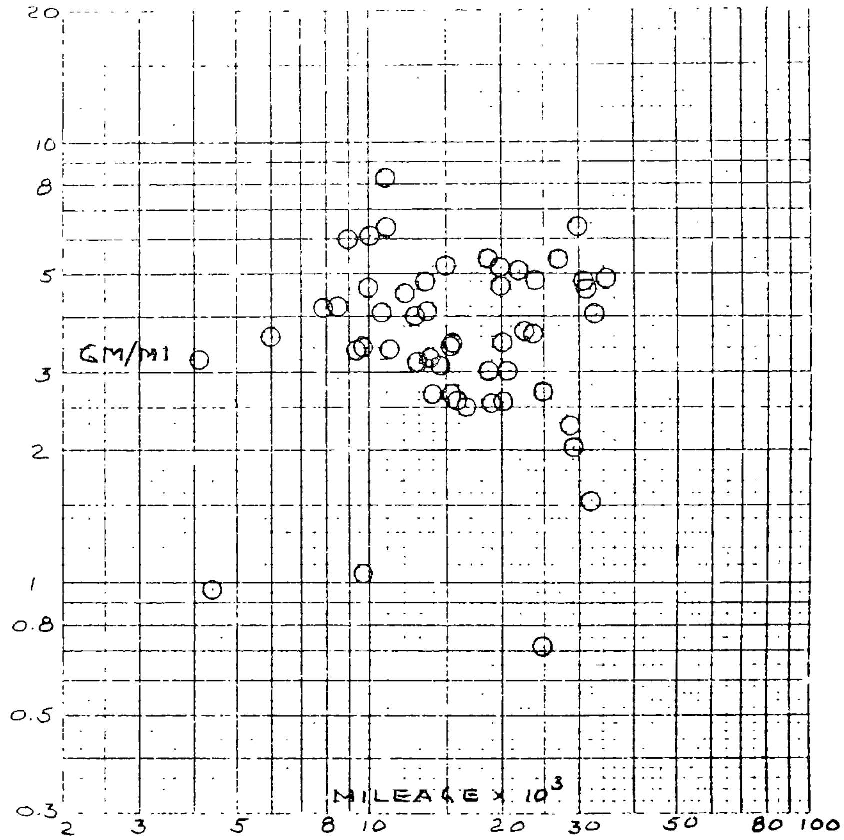


Figure 12 - NO_2 Emissions of Hot-Stable Bag 4 Versus Accumulated Mileage

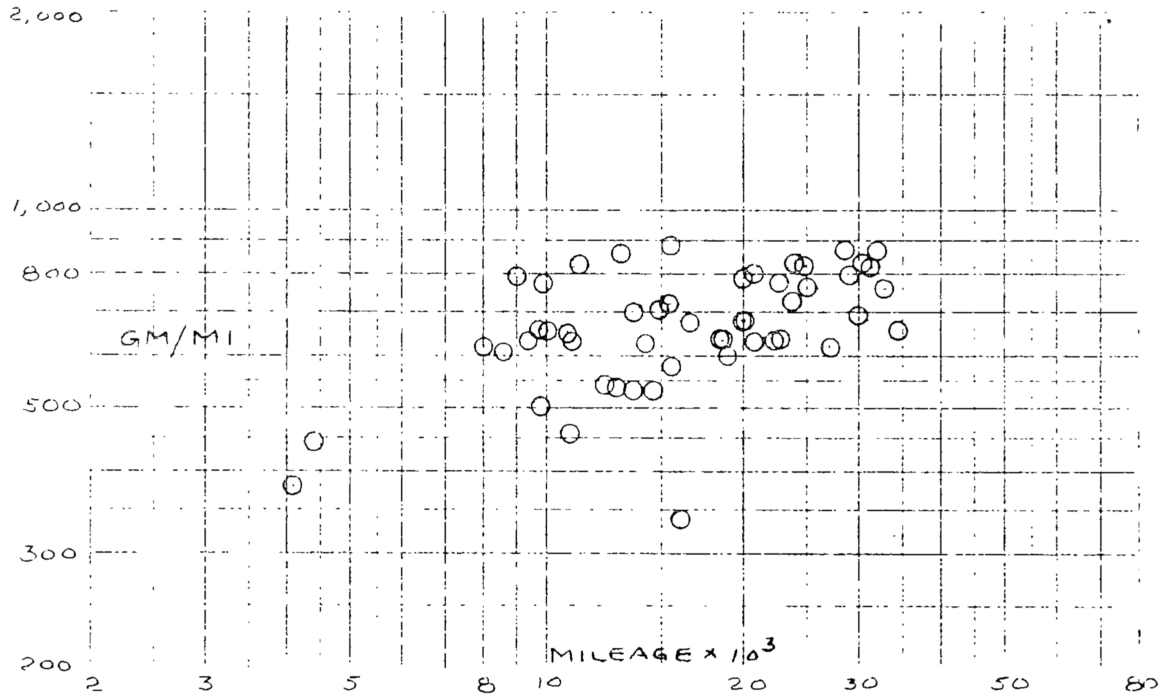


Figure 13 - CO_2 Emissions of Cold-Transient Bag 1 Versus Accumulated Mileage

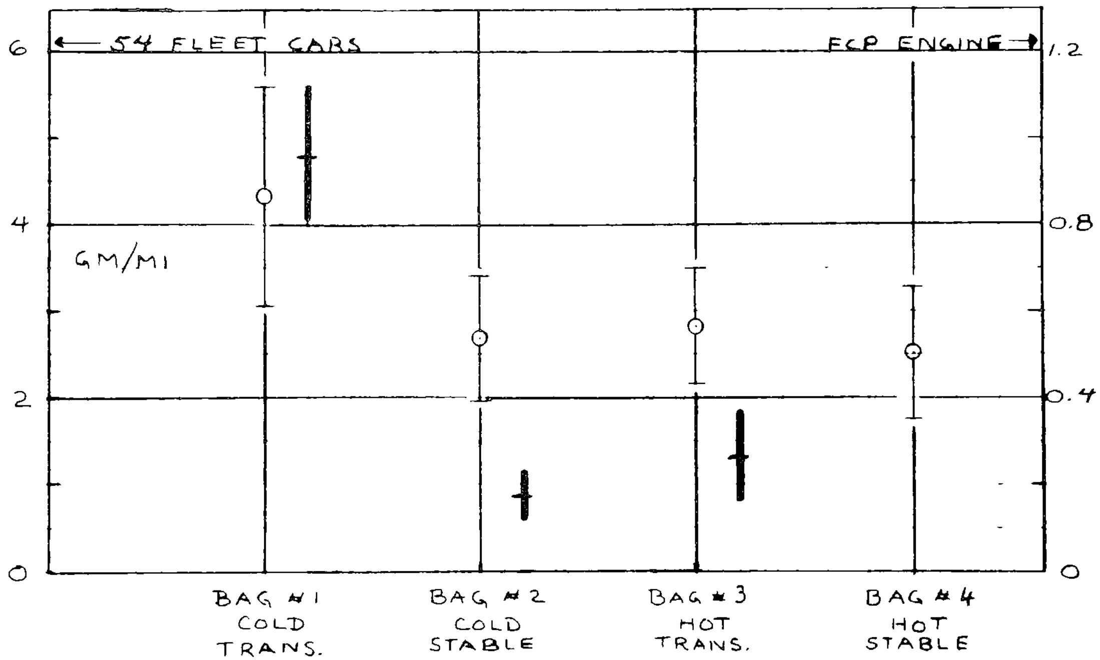


Figure 14 - Mean and Standard Deviation of HC

54 Fleet Cars

Jeep with FCP Engine (EPA)

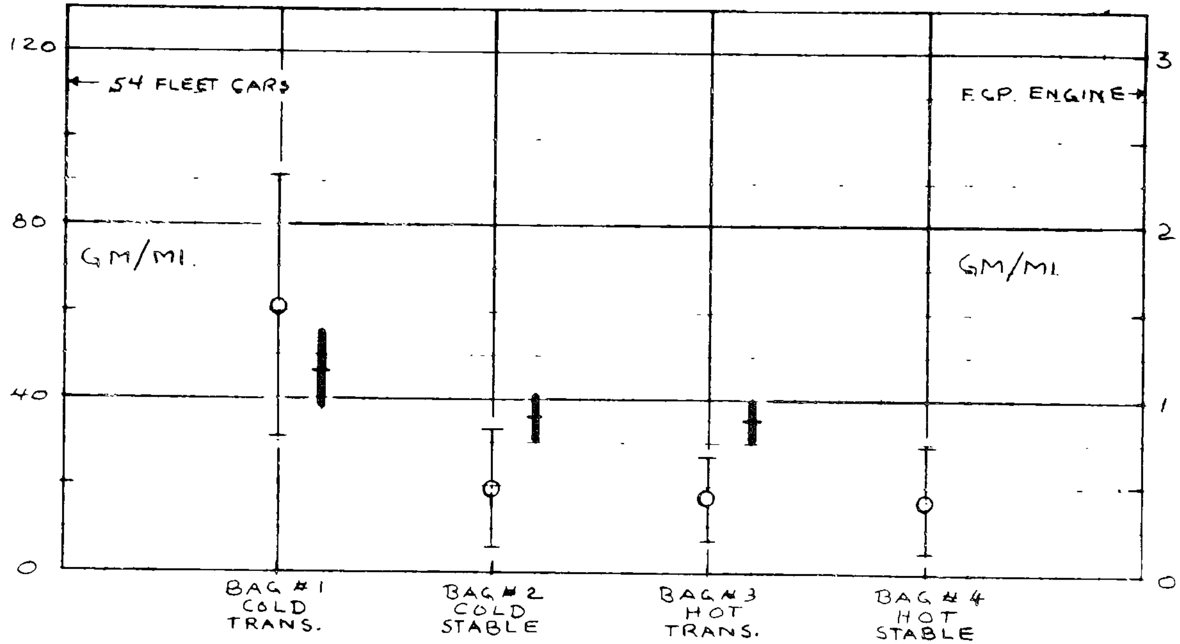


Figure 15 - Mean and Standard Deviation of CO

54 Fleet Cars

Jeep with FCP Engine (EPA)

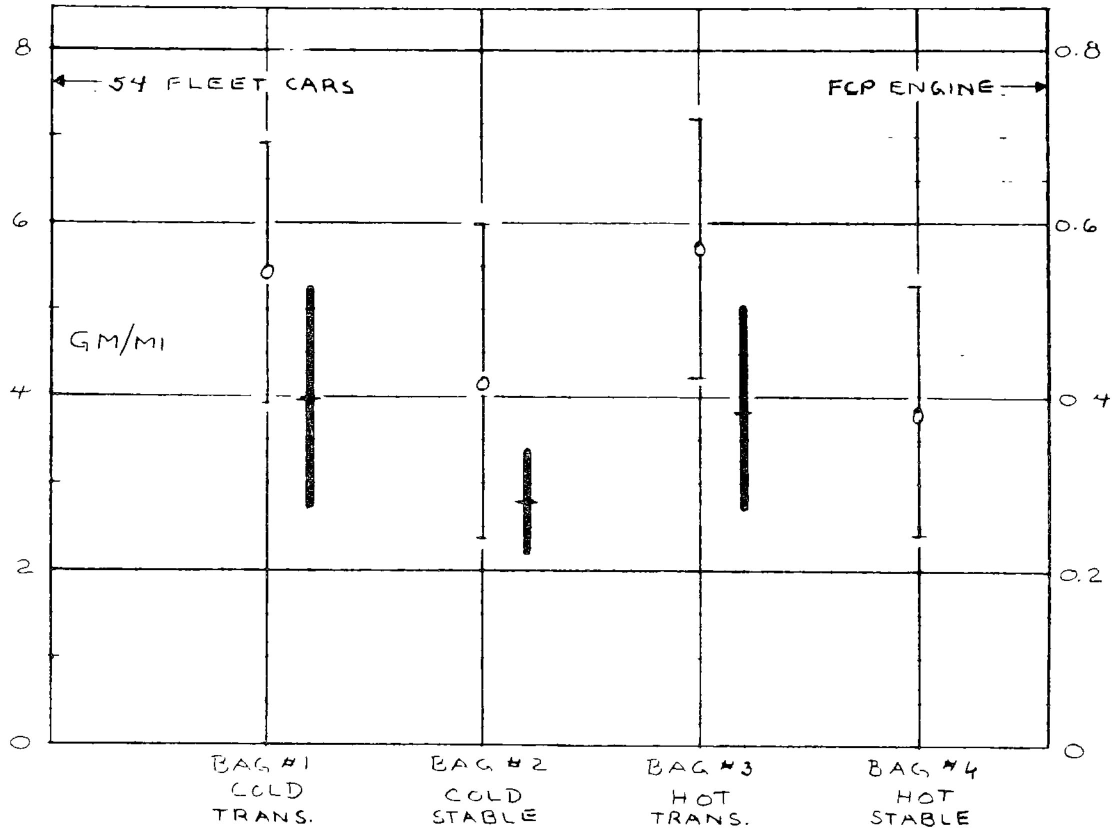


Figure 16 - Mean and Standard Deviation of NO_x

54 Fleet Cars

Jeep with FCP Engine (EPA)

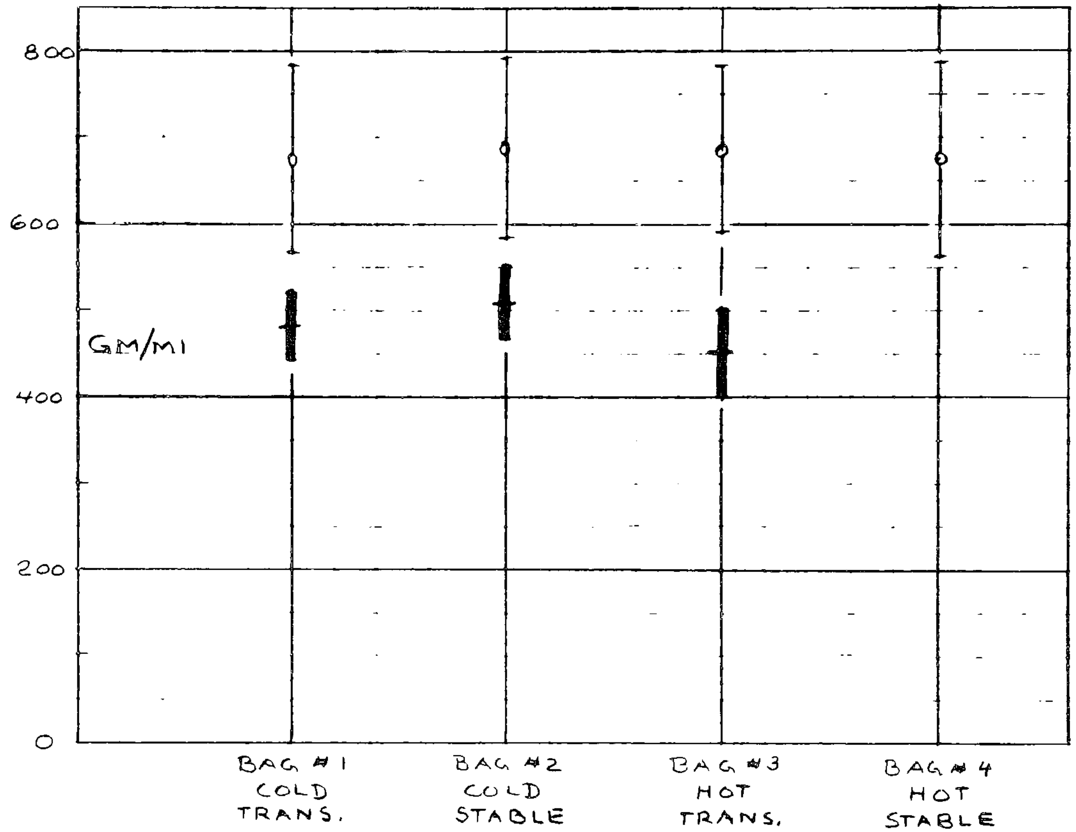


Figure 17 - Mean and Standard Deviation of CO_2

54 Fleet Cars

Jeep with FCP Engine (EPA)

APPENDIX F EXHAUST EMISSIONS FOR CHEVROLET AND FORD AUTOMOBILES

H. T. McAdams

Data on five Chevrolet and five Ford automobiles have been made available for analysis of the effect of mileage accumulation on exhaust emissions. The Chevrolets were 1969 Impala sedans and the Fords were 1969 Galaxie sedans. The data were reported by Automotive Research Associates, Inc. under Contract CPA-22-69-140 for the Environmental Protection Agency.* Though these tests failed to show a linear relationship between mileage and emission of HC and CO, they can be used to formulate an analysis of variance in which "within-automobile" sources of variance can be assessed in relation to "between-automobile" sources of variance.

The term "within-automobile variance" here is interpreted to mean those differences in emissions which are observed when the same automobile is tested for emissions at different times. These successive tests represent different mileage accumulations but, inasmuch as regression analysis failed to show a significant trend with mileage, it is permissible to conclude that the successive tests exhibit differences attributable to random causes. These differences are due, in part, to testing errors but may reflect actual changes in vehicle performance from test to test. The term "between-automobile variance" is interpreted to mean those differances in emissions which are associated with the individual idiosyncrasies of the several presumably identical cars tested.

By comparing the magnitudes of the "between" and "within" variance, one can estimate the effect which individual differences among cars might be expected to contribute to emission assessment. These differences presumably arise, at least in part, from manufacturing tolerances. Thus it was felt that by separating "between" from "within" variability, one might be able to speculate on the effect of manufacturing tolerances on stratified charge emissions.

Analysis of variance was performed by considering variation within cars as being nested within car-to-car variation. Thus, in a certain sense, the several mileage checks on a particular car were considered as replicates or repeat tests of that car. It was naturally reasoned that this variation

---

* Relationship of Engine Deterioration to Exhaust Emissions. Final Report. May 28, 1971. Automotive Research Associates, Inc. San Antonio, Texas.

would be less than if one looked at the several mileage points without noting the identity of the vehicle. Only that subset of mileage accumulations were used which were common to all five vehicles of a particular make (i.e., either Chevy or Ford).

Each group of vehicles tested exhibited seven mileage points which were common to all vehicles in the group. Though the seven mileage points for the Chevy vehicles were not the same as the seven mileage points for the Ford vehicles, it was felt that both provided a fair assessment of within-vehicle variation. For each make of automobile, therefore, there were 4 degrees of freedom for assessing between-vehicle variance and 6 x 5 30 degrees of freedom for assessing within-vehicle variance.

A convenient measure of the relative importance of these two sources of variability is provided by a statistic called the intraclass correlation coefficient. It can be shown that the usual "mean squares within", as computed from analysis of variance, is a measure of σ^2, the "within" variance. The "between mean squares", however, is an estimate of $\sigma^2 + 7\sigma_a^2$, where σ_a^2 is the "between-automobile" component of variance. One can thus formulate two equations

$$\sigma^2 + 7\sigma_a^2 = \text{mean squares among}$$

$$\sigma^2 - \text{mean squares within}$$

which can be solved for σ^2 and σ_a^2. Then $\sigma^2 + \sigma_a^2$ is an estimate of what might be termed the "combined variance." It is the variance which would be observed if each automobile were tested once only and the standard deviation of results was computed from these tests. The ratio

$$\frac{\sigma_a^2}{\sigma^2 + \sigma_a^2}$$

is the fraction of this combined variance which is accounted for by vehicle-to-vehicle differences.

The data to be analyzed were regarded as consisting of 12 measures. The first six denote HC emissions, the last six CO emissions. In each group of six measures, there were hot cycle, cold cycle and combined tests, each test being evaluated in terms of constant volume sampling (CVS) and concentration (conc) measurements.

Results of analysis of variance on the Chevy and Ford automobiles are provided in Tables I and II respectively. Two sets of F-ratios were computed. Those tabulated in the rightmost column of each table are intended to answer the question of whether the between-variance is significant for the particular make of vehicle - i.e., Chevrolet or Ford. If the F-ratio is significant at the 0.05 level, it is labeled with an asterisk. The F-ratios tabulated in the "Mean Square between" and "Mean Square within" columns provide a measure of whether significant differences exist between the variability of Ford and Chevrolet cars. If it is noted that the F-ratio for "MS within" is significant, this fact implies that one make of automobile shows significantly more variability in successive tests of the same automobile than does the other make. On the other hand if the F-ratio for "MS between" is significant, it is indicated that one make did not reproduce as well as the other in the vehicle-to-vehicle sense. This fact could imply that manufacturing variability for one make is greater than for the other. If one of the F-ratios discussed here is significant, it is followed by either a (C) or an (F). The presence of the letter denotes significances at the 0.05 level; C or F denotes, respectively, whether the Chevrolets or Fords exhibited the higher mean squares.

In general, it appears that the Chevrolets were more variable than the Fords. This fact is further evidenced by the magnitude of the intraclass correlation coefficients. Note that these range from a few percent up to more than 50%.

Tables III and IV were constructed by first transforming the emission measurement values to their natural logarithms. These logarithms were then subjected to analysis of variance as before. Differences between the Chevy and Ford vehicles, especially as to their CO emissions, are now more pronounced.

Justification for performing analysis of variance on logarithmically transformed data resides in the fact that variability in emissions tends to be proportional to the mean level of emissions. It can be shown that the taking of logarithms provides a "variance-stabilizing transformation" so that variability of the transformed data is substantially independent of its mean magnitude.

On the basis of these tests it is concluded that appreciable variation in emissions can exist among vehicles of the same make. Thus it can not be concluded that engine-to-engine variability is negligible. On the other hand, by virtue of the fact that one make of automobile in these tests was less variable than the other, the prospect for good reproducibility of engines is offered. Which of these circumstances is most applicable to stratified-charge engines, or whether either is, is unknown.

TABLE 1

HC EMISSIONS FOR CHEVY AND FORD AUTOMOBILES

	Mean	Std.Dev.	Coef.Var.	ANALYSIS OF VARIANCE			
				MS Between	MS Within	Intra Cor	F
Measure 1 HC-Cold-CVS							
Chevy	13.138	17.745	1.351	430.551	295.593	0.061	1.457
Ford	15.070	10.088	0.669	135.715	96.110	0.056	1.412
F-Ratio				2.951	3.076		
Measure 2 HC-Cold-Conc							
Chevy	4.816	1.273	0.264	3.559	1.299	0.199	2.741*
Ford	5.690	1.710	0.300	5.302	2.527	0.136	2.098
F-Ratio				1.490	1.945		
Measure 3 HC-Hot-CVS							
Chevy	10,486	4.339	0.414	21.213	18.429	0.021	1.151
Ford	9.620	2.822	0.293	10.432	7.552	0.052	1.381
F-Ratio				2.033	2.440		
Measure 4 HC-Hot-Conc.							
Chevy	4.911	1.576	0.321	3.888	2.252	0.094	1.727
Ford	4.462	1.485	0.333	4.178	1.877	0.149	2.225
F-Ratio				1.075	1.200		
Measure 5 HC-CVS-Composite							
Chevy	10.315	5.162	0.500	61.042	20.910	0.215	2.919*
Ford	11.830	4.424	0.374	19.245	19.624	-0.003	0.981
F-Ratio				3.172	1.066		
Measure 6 HC-Conc-Composite							
Chevy	4.757	1.261	0.265	1.859	1.544	0.028	1.204
Ford	5.010	1.380	0.276	3.932	1.568	0.177	2.509
F-Ratio				2.115	1.016		

* Denotes significance at 0.05 level

TABLE II

CO EMISSIONS FOR CHEVY AND FORD AUTOMOBILES

	Mean	Std.Dev.	Coef.Var.	ANALYSIS OF VARIANCE MS Between	MS Within	Intra Cor	F
Measure 7 CO-Cold-CVS							
Chevy	41.435	26.392	0.637	2902.709	328.871	0.528	8.826*
Ford	64.558	22.034	0.341	1128.609	378.287	0.221	2.983*
F-Ratio				2.572	1.150		
Measure 8 CO-Cold-Conc							
Chevy	35.825	27.982	0.781	3038.193	407.120	0.480	7.463*
Ford	42.247	16.622	0.393	190.921	290.528	-0.052	0.657
F-Ratio				15.913(C)	1.401		
Measure 9 CO-Hot-CVS							
Chevy	20.917	13.562	0.648	806.742	80.112	0.564	10.070*
Ford	12.889	7.665	0.595	93.230	53.001	0.098	1.759
F-Ratio				8.653(C)	1.512		
Measure 10 CO-Hot-Conc							
Chevy	23.909	19.227	0.804	2157.386	71.728	0.806	30.077*
Ford	9.880	9.282	0.939	144.393	76.449	0.113	1.889
F-Ratio				14.941(C)	1.066		
Measure 11 CO-CVS-Composite							
Chevy	37.239	39.677	1.065	4518.391	1083.540	0.312	4.170*
Ford	38.441	13.081	0.340	263.245	155.750	0.090	1.690
F-Ratio				17.164(C)	6.957(C)		
Measure 12 CO-Conc-Composite							
Chevy	28.195	20.116	0.713	2141.301	115.194	0.715	18.589*
Ford	20.441	5.733	0.280	69.427	26.779	0.185	2.593
F-Ratio				30.842(C)	4.302(C)		

* Denotes significance at 0.05 level

TABLE III

LOGARITHMICALLY TRANSFORMED HC EMISSIONS FOR CHEVY AND FORD AUTOMOBILES

	Mean	Std.Dev.	Coef.Var.	ANALYSIS OF VARIANCE MS Between	MS Within	Intra Cor	F
Measure 1 HC-Cold-CVS							
Chevy	2.265	0.653	0.388	1.081	0.317	0.256	3.414
Ford	2.578	0.478	0.185	0.250	0.225	0.015	1.110
F-Ratio				4.324	1.409		
Measure 2 HC-Cold-Conc							
Chevy	1.542	0.247	0.160	0.116	0.052	0.151	2.248
Ford	1.669	0.454	0.272	0.258	0.198	0.041	1.301
F-Ratio				2.224	3.808		
Measure 3 HC-Hot-CVS							
Chevy	2.289	0.339	0.148	0.109	0.116	-0.008	0.944
Ford	2.218	0.322	0.145	0.141	0.098	0.60	1.445
F-Ratio				1.294	1.184		
Measure 4 HC-Hot-Conc							
Chevy	1.547	0.298	0.193	0.139	0.080	0.094	1.726
Ford	1.441	0.342	0.237	0.154	0.111	0.052	1.387
F-Ratio				1.108	1.388		
Measure 5 HC-CVS-Composite							
Chevy	2.256	0.382	0.169	0.367	0.109	0.254	3.383
Ford	2.415	0.330	0.137	0.105	0.109	-0.005	0.962
F-Ratio				3.495	1.000		
Measure 6 HC-Conc-Composite							
Chevy	1.528	0.252	0.165	0.072	0.062	0.022	1.154
Ford	1.574	0.286	0.182	0.112	0.077	0.063	1.467
F-Ratio				1.556	1.242		

* Denotes significance at 0.05 level

TABLE IV

LOGARITHMICALLY TRANSFORMED CO EMISSIONS FOR CHEVY AND FORD AUTOMOBILES

	Mean	Std.Dev.	Coef.Var.	ANALYSIS OF VARIANCE			
				MS BETWEEN	MS Within	Intra Cor	F
Measure 7 CO-Cold-CVS							
Chevy	3.520	0.744	0.211	2.466	0.235	0.576	10.504*
Ford	4.089	0.464	0.114	0.389	0.187	0.134	2.084
F-Ratio				6.339	1.257		
Measure 8 CO-Cold-Conc							
Chevy	3.386	0.632	0.187	1.571	0.205	0.488	7.666*
Ford	3.672	0.393	0.107	1.102	0.163	-0.056	0.627
F-Ratio				1.540	1.258		
Measure 9 CO-H T-Cus							
Chevy	2.848	0.681	0.239	1.924	0.220	0.525	8.747*
Ford	2.424	0.510	0.210	0.345	0.246	0.054	1.401
F-Ratio				5.577	1.118		
Measure 10 CO-Hot-Conc							
Chevy	2.870	0.921	0.321	4.762	0.196	0.769	24.297*
Ford	2.071	0.625	0.302	0.893	0.306	0.215	2.915*
F-Ratio				5.333	1.561		
Measure 11 CO-CVS-Composite							
Chevy	3.351	0.729	0.217	2.118	0.267	0.498	7.945*
Ford	3.579	0.416	0.116	0.225	0.164	0.050	1.368
F-Ratio				9.413(C)	1.628		
Measure 12 CO-Conc-Composite							
Chevy	3.093	0.886	0.286	2.431	0.510	0.350	4.768*
Ford	2.980	0.285	0.096	0.167	0.067	0.176	2.500
F-Ratio				14.557(C)	7.612(C)		

* Denotes significance at 0.05 level

APPENDIX G - THE LOG NORMAL DISTRIBUTION AS A STATISTICAL MODEL FOR EXHAUST EMISSIONS

H. T. McAdams

1. INTRODUCTION

The log-normal distribution has been indicated, empirically, to be applicable to the statistical distribution of exhaust emissions. A foundation for this observation is sought in order to substantiate the use of this model on physical grounds. First, the distribution and its properties are examined and, second, assumptions which might give rise to the distribution in an emissions context are explored.

2. PROPERTIES OF THE LOG-NORMAL DISTRIBUTION

A random variable X is said to have a log-normal distribution if log X has a normal distribution. Let Y = log X. Then $X = e^Y$ and we want to find the expected value and variance of e^Y.

Consider

$$E[X] = E[e^Y] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{y-m}{\sigma}\right)^2} \cdot e^y \, dy \tag{1}$$

where E denotes expectation, and note that the moment-generating function of the random variable Y is

$$E[e^{ty}] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{y-m}{\sigma}\right)^2} \cdot e^{ty} = e^{tm+\frac{1}{2}t^2\sigma^2} \tag{2}$$

Now when t = 1, (2) is formally equivalent to (1), so that

$$E[X] = e^{m+\frac{1}{2}\sigma^2} \tag{3}$$

where m is what is sometimes called (incorrectly) the "log population mean" and σ is what is similarly called "log population standard deviation."

Similarly, by definition of the variance of a random variable

$$Var[X] = E[X^2] - \{E[X]\}^2$$

But

$$E[X^2] = E[e^{2Y}] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{y-m}{\sigma}\right)^2} \cdot e^{2y}\, dy \qquad (4)$$

which is formally equivalent to (2) when t = 2.
Therefore,

$$E[X^2] = e^{2m+2\sigma^2}$$

and

$$\{E[X]\}^2 = e^{2m+\sigma^2}$$

so that

$$Var[X] = e^{2m+2\sigma^2} - e^{2m+\sigma^2} \qquad (5)$$

Manipulation of the results of (3) and (5) gives rise to some interesting relationships between the mean and standard deviation of the log-normal distribution and its percentile points. Let us, for convenience, rewrite (3) and (5) as

$$\bar{x} = e^{m+\frac{1}{2}\sigma^2} \qquad (6)$$

and

$$s^2 = e^{2m+2\sigma^2} - e^{2m+\sigma^2} \tag{7}$$

where $\bar{x}$ denotes the mean and s denotes the standard deviation of the random variable X. Then

$$\bar{x}^2 + s^2 = e^{2m+\sigma^2} + e^{2m+2\sigma^2} - e^{2m+\sigma^2}$$

$$= e^{2m+2\sigma^2}$$

or

$$\bar{x}^2 + s^2 = e^{\sigma^2} \cdot e^{2m+\sigma^2}$$

$$= \bar{x}^2 e^{\sigma^2}$$

Then $$e^{\sigma^2} = \frac{\bar{x}^2 + s^2}{\bar{x}^2} \quad \text{oR} \quad e^{\frac{1}{2}\sigma^2} = \frac{\sqrt{\bar{x}^2 + s^2}}{\bar{x}} \tag{8}$$

Also, from (6),

$$e^m = \frac{\bar{x}}{e^{1/2\,\sigma^2}} \tag{9}$$

Combining (8) and (9), one obtains

$$e^m = \frac{\bar{x}}{\frac{\sqrt{\bar{x}^2 + s^2}}{\bar{x}}} = \frac{\bar{x}^2}{\sqrt{\bar{x}^2 + s^2}} \tag{10}$$

But m is the mean logarithm and, since the logarithms are normally distributed m is also the median logarithm. If 50% of the logarithms are less than m, then 50% of the antilogarithms are less than e^m -- in short, e^m is the median x_{50} of the log-normal distribution. Therefore,

$$x_{50} = \frac{\bar{x}^2}{\sqrt{\bar{x}^2 + s^2}} \tag{11}$$

Let us now consider the 84th percentile x_{84}. In the log-transform space, this is the value $m + \sigma$, or, in antilog space,

$$e^{m+\sigma} = e^{m} \cdot e^{\sigma}$$

$$= e^{\sigma}\, x_{50} \qquad (11)$$

From (8) we have, taking logarithms of both sides,

$$\sigma^2 \quad \ln\left(\frac{\bar{x}^2 + s^2}{\bar{x}^2}\right)$$

$$\sigma = \sqrt{\ln\left(\frac{\bar{x}^2 + s^2}{\bar{x}^2}\right)}$$

or

$$e^{\sigma} = e^{\sqrt{\ln\left(\frac{\bar{x}^2 + s^2}{\bar{x}^2}\right)}} \qquad (12)$$

Then, combining (11) and (12), one has

$$x_{84} \quad e^{m+\sigma} = x_{50}\, e^{\sqrt{\ln\left(\frac{\bar{x}^2 + s^2}{\bar{x}^2}\right)}}$$

By similar reasoning it can be shown that

$$x_{16} = e^{m-\sigma} = \frac{x_{50}}{e^{\sqrt{\ln\left(\frac{\bar{x}^2 + s^2}{\bar{x}^2}\right)}}} \qquad (13)$$

or

$$x_{16} = \left(\frac{x_{50}}{x_{84}}\right) \cdot x_{50}$$

This suggests that percentiles can be computed as simple multiples of the median.

Further generalization is, in fact, straightforward. For, consider a real number k and the quantity $m + k\sigma$. From tables of the cumulative distribution function for a normal distribution, a value of k can be determined for any desired percentile. Then, in antilog space,

$$e^{m+k\sigma} = e^{m} \cdot e^{k\sigma}$$

defines exactly the same percentile

or

$$e^{m+k\sigma} = X_{50} \cdot e^{k\sigma}$$

$$= X_{50} \cdot \left(e^{\sigma}\right)^{k}$$

Thus it is seen that the antilog of σ, the log standard deviation, provides a multiplying factor applicable to the median in exactly the same way that an additive factor is applied to the mean of a normal distribution. In practice, computations are simpler if the quantities m and σ are computed by transforming observations to their logarithmic equivalents.

3. ORIGIN OF THE LOG-NORMAL DISTRIBUTION

If emissions measurements tend to follow a log-normal distribution, it is of interest to ask why this relationship holds. If a logical basis for the log-normal distribution can be established, considerable light may be shed on the nature of error propagation, the requirements to be satisfied in complying with emission standards, hardware requirements needed to answer questions of technical feasibility, and the form that test plans should assume.

A statistical process may tend toward producing a log-normal distribution if the following conditions are met:

a) Several sources of error or variability combine in generating the random variable under study.

b) These sources combine multiplicatively rather than additively.

Within reasonable limits, the form of the statistical distribution of the source variables is irrelevant.

To justify the above hypothesis, at least heuristically, consider the central limit theorem of mathematical statistics. Though this theorem can be expressed in various ways, the following statement suffices for present purposes. If an arbitrary population distribution has mean μ and finite variance σ^2, then the distribution of the sample mean approaches the normal distribution with mean μ and variance σ^2/n as the sample size n increases. Sometimes it is said that the distribution of the sample means tends asymtotically to a normal distribution. How large n has to be in order to approach closely to a normal distribution depends on the form of the distribution of the variables being combined. As has been shown in experience with statistical quality control concepts, normality can be closely approached if n is as small as 3 or 4, even if the distributions of the original distributions depart markedly from a normal distribution.

Now consider the case of n sources of error--that is, consider n random variables $X_1, X_2, \ldots, X_n$. Perform the logarithmic transformation

$$\log X \quad \log X_1 + \text{og } X_2 + \ldots + \log Xn$$

or

$$Y = Y_1 + Y_2 + \ldots + Y_n$$

where $Y_i = \log X_i$

Then, if the central limit theorem applies, Y will tend asymmtotically toward a normal distribution, and this fact implies that in antilog space X will tend toward a log-normal distribution.

Is there any reason to believe that the sources of variability in automotive exhaust emissions should combine multiplicatively? Perhaps there is. Consider, first of all, errors of measurement in determining the concentration of a particular pollutant in the exhaust stream. Let the error involved in measuring the concentration C be measured by its standard deviation σ_c. There is reason to believe that σ_c is proportional to C--that is, that the relative standard deviation or coefficient of variation σ_c/C is constant for a particular effluent and a particular measurement situation.

Now, consider a particular engine and let its emissions be measured repeatedly by the process under consideration. The actual concentration C will vary from time to time, due to differences in ambient environmental conditions, condition of linkages and engine adjustments, type of fuel, operator-induced variations and the like. It will follow, therefore, that σ_c will also vary proportionally to C and that the net result will be analogous to the multiplication of two sources of error.

Let us further consider the relationship which obtains between repeated tests of the same automobile and variation among many automobiles in a fleet or in a production output. It might, perhaps, be argued that if a particular automobile has a high mean level of emissions, that same automobile may tend to have highly-variable emissions. Conversely, an automobile having a relatively low mean level of emissions may tend to have relatively constant emissions. A moment's reflection will show that these statements must, to some extent at least, be true. Since emissions are bounded on the low side of the scale (they can not be less than zero), it is impossible for individual emission measurements to range upward beyond a certain level and yet maintain a low mean, unless the high measurements are offset by a large number of measurements close to zero. This bounding process will tend to limit the magnitude of the standard deviation and, at the same time, may tend to produce a skewed distribution which, itself, may resemble a log-normal distribution.

Now consider the combination of vehicle-to-vehicle variability, time-to-time variability within a given vehicle, and test-to-test variability at a given time. If test-to-test variability is proportional to the mean level of the particular automobile at a given time, and time-to-time variability is proportional to the mean level for the particular vehicle, then the three sources of variability would combine in a multiplicatively way. The result, by virtue of the central limit theorem, would be a log-normal distribution.

4. IMPLICATIONS OF STATISTICAL TREATMENT OF EMISSION MEASUREMENTS

The implications for statistical analysis of emission measurements is clear. Whereas in normally distributed data our concern is with two quantities, the mean and standard deviation of the measurements, here we are concerned with m and σ, the mean and standard deviation of the logarithms of the measurements. If one reasons in antilog space, the corresponding quantities are

$$e^{m} = \text{the median of the measurements}$$

and

$$e^{\sigma} = \text{the "ratio standard deviation".}$$

The quantity e^{σ} is a fundamental quantity which defines the dispersion of the results and which, together with the median, can be used to define any percentile of the distribution. Thus the median and the ratio standard deviation play the role that the mean and simple standard deviation usually play. Note that the use of a ratio standard deviation applied to the median does not differ radically from the notion of a constant coefficient of variation applied to the mean.

If is of interest, in fact, to compare the ratio standard deviation with the coefficient of variation and to provide a means for converting from one to the other. The coefficient of variation is

$$\frac{s}{\bar{x}} = \frac{\sqrt{e^{2m+2\sigma^2} - e^{2m+\sigma^2}}}{e^{m+1/2\sigma^2}}$$

$$= \sqrt{\frac{e^{2m+2\sigma^2} - e^{2m+\sigma^2}}{e^{2m+\sigma^2}}}$$

$$= \sqrt{e^{\sigma^2} - 1} \qquad (14)$$

A more usable approximation is provided by expanding e^{σ^2} to an acceptable number of terms:

$$e^{\sigma^2} = 1 + \sigma^2 + \sigma^4/2 + \ldots$$

For small σ^2, it suffices to retain only the first two terms, so that

$$\frac{s}{\bar{x}} \approx \sqrt{1+\sigma^2-1} = \sigma$$

Then

$$e^{\sigma} = e^{s/\bar{x}} \approx 1 + s/\bar{x}$$

or

Ratio standard deviation $\approx$ 1 + coefficient of variation

For example, if the coefficient of variation is 0.2, the ratio standard deviation is approximately 1.2. An exact computation according to (14) would yield

$$e^{\sigma^2} \quad 1 = (0.2)^2$$

$$e^{\sigma^2} = 1.04$$

$$\sigma^2 \quad \ln 1.04 = 0.03922$$

$$\sigma = .192$$

or

$$e^{\sigma} = 1.212$$

Thus, for approximation, the ratio standard deviation can be estimated by adding 1 to the coefficient of variation.

Further implication of the analysis is that, in analysis of variance aimed at estimating sources of variability, the emissions data can be subjected first to logarithmic transformation. Analysis of variance performed in log space would yield components of variance combining multiplicatively, according to the postulated basis of the log-normal distribution of emissions.

APPENDIX H EXPERIMENT DESIGN CONSIDERATIONS FOR EMISSIONS TESTING OF STRATIFIED—CHARGE ENGINES

H. T McAdams

1. INTRODUCTION

A test program to evaluate emissions for stratified charge engines must take into consideration a number of variables which can affect the results. Variables which have been recognized by EPA as being important to emissions are listed in Table I.

TABLE I

Test Variables

- Engine Supplier (Ford/Texaco)
- Transmission/Drive (4 Wheel Stick/2 Wheel Auto)
- Engine Tuning (Max.Fuel Economy/Min.Emissions)
- Catalyst (None/Platinum/Palladium)
- EGR (None/A Little/A Lot)
- Aspiration (Natural/Turbo)
- Fuel (Gasoline/CITE/Diesel)
- Mileage
- Inertia
- Road Load
- No. Replications
- No. Operators

Annotation of these variables, as well as the postulation and discussion of other sources of variability, is considered to be in order before proceeding to a formulation of test programs for either available hardware or for hardware to be procured and tested in a later time frame.

In previous presentations and documents, it has been shown that two major sources of variability affect emissions as determined from an emissions test. In one source are all those factors which affect the measurement process itself: the test driving cycle, the operator, the instrumentation accuracy and precision, the ambient environmental conditions (temperature, humidity, atmospheric pressure). In the other source are all those factors which affect the engine and/or its performance: its basic design (e.g., whether Ford SCE, Texaco SCE, conventional, etc.), the transmission or drive mechanism to which it is coupled, the load to which it is subjected, the age and condition of the engine, the type of fuel used, the manner in which the engine is tuned, and the number and type of emission-control devices employed. In the light of the many sources of variability, which the above catalog by no means exhausts, it becomes very difficult if not impossible to answer the general question, "What numbers, in terms of - say - grams pollutant per mile, characterize the stratified charge engine?" The question clearly has little meaning until it has been adequately qualified by specification of such key items as type of test employed, fuel used, and so on. On the other hand, complete particularization of all the factors which can affect emission measurements is neither possible nor desirable. Rather, certain of these effects are more appropriately addressed by randomization and statistical aggregation. An important decision prior to the formulation of a testing program is to decide which variables are to be controlled and which are to be randomized.

The distinction between fixed and random effects can be better appreciated by considering the following two mathematical formulations or models.

MODEL I:
$$x_{ij} = \mu + \alpha_j + \epsilon_{ij}$$
α_j fixed effect for the j^{th} column
$$\epsilon_{ij} \sim N(0, \sigma^2)$$

MODEL II:
$$x_{ij} = \mu + a_j + \epsilon_{ij}$$
$$a_j \sim N(0, \sigma_A^2)$$
$$\epsilon_{ij} \sim N(0, \sigma^2)$$

In these models, it is assumed that data subject to two-way classification is available and that this data has been arranged in rows (indexed by i) and columns (indexed by j). In Model I, the "fixed-effects" model, it is assumed that all data belonging to the j^{th} column contributes a fixed additive quantity α_j to the value x_{ij} listed in the i^{th} row of the j^{th} column. Furthermore, it is assumed that it is inappropriate to consider the j^{th} constraint, whatever it might be, as belonging to a statistical population. In short, the j^{th} constraint is not imposed by chance but by intelligence and control and is repeatable with negligible error. In Model II, the "random-effects" model, it is assumed that the differences which exist among columns are induced by random rather than by conscious choice and that there is no direct way by which the contribution a_j can be exactly reproduced in a subsequent experiment. In short, a_j is an outcome or realization of a random variable A assumed to be normally distributed with mean 0 and variance σ_A^2. It is meaningless to attempt to reproduce the results for say column 3, because either there is no available basis for choice or the basis for choice has been relinquished in the interest of simplifying the problem under consideration This being the case, a new choice for column 3 will differ from the original choice in the same way as do the other columns in the array.

An example of each of these models in the context of emissions testing will help to clarify their implications.

Suppose, for example, that a particular vehicle has been tested several times at each of three inertial weights: 2000 pounds, 2500 pounds and 3000 pounds. These three sets of tests constitute the columns, and the entries within the columns constitute replicate test runs. Variation within the columns result from random causes affecting the testing and measurement operation and can be considered to be normally distributed with mean 0 and variance σ^2. The contribution α_1 of the first column, however, is peculiar to an inertial weight of 2000 pounds. If the test is abandoned and later returned to, it should be possible to obtain virtually the same α_1 simply by setting up the experiment so as to provide an inertial weight of 2000 pounds. Furthermore, it makes no sense to consider 2000 pounds as the outcome of a random variable so long as it is practicable to treat it as an assignable cause under the control of the experimenter.

Consider, now, a situation in which inertial weight is fixed at a particular value- say 2500 pounds but three vehicles are obtained from the manufacturer for test. There is nothing about the three vehicles that distinguishes one from the other or from other vehicles which might still be in the manufacturer's inventory. Nevertheless, upon being tested, the vehicles exhibit distinct differences in emissions, as observed by differences among the column means. Though a value a_j is uniquely ascribable to each of the three vehicles, it is meaningless to attempt to select a fourth vehicle duplicating either a_1, a_2, or a_3 because there is no basis on which such a selection can be made. Nevertheless, we can assume that the variation among the $a_j (j = 1,2,3)$ reflects a stable property of subsequent choices of vehicles, in that it provides an estimate of σ_A^2, the variance of the statistical distribution associated with the vehicle population.

It is hurriedly pointed out that the above examples are oversimplified and that the distinction between fixed and random effects is somewhat chimerical. In the real world, vehicles may well exhibit a spectrum of inertial weights and it may be appropriate to characterize these weight variations as a statistical distribution. The weight variations could accordingly induce a statistical distribution of emissions, as shown in Figure 1.

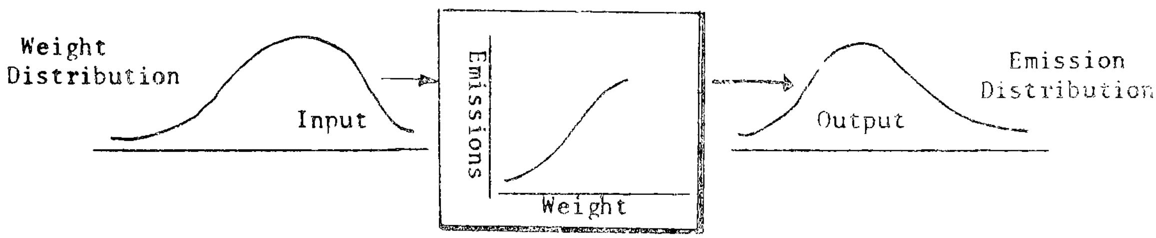


FIGURE 1

RANDOM EFFECTS AS AN INDUCED DISTRIBUTION

Testing procedures and the attendant data processing will differ appreciably depending on whether it is desired to qualify vehicles at a fixed weight or to qualify vehicles collectively according to the prevailing frequency distribution of weights. Similar questions arise in the structuring of other variables which affect emissions measurements, either through the physical attributes of the vehicle and its performance capabilities or through variables affecting the test operation. For example, vehicle operators or inter-laboratory differences in instrumentation may well contribute differences in emissions measurements. Is it better to attempt to "standardize" these influences or to aggregate them statistically and allow their influence to be felt as statistical "errors"?

Though it is a point of some degree of subtlety, it will be evident that the emissions distribution arising from Figure 1 differs in an important respect from the distribution of errors ϵ_{ij} in models I and II. It is axiomatic that the ϵ_{ij} are induced in a manner analogous to Figure 1, but in the case of the ϵ_{ij} the underlying distribution of causes is unknown and not determinable. Consequently, if we are to estimate the distribution of the ϵ_{ij}, we must measure the distribution at the output. On the other hand, there are variables such as inertial weight which can be varied parametrically so as to determine the input-output relationship. Once this relationship is known, it can be combined with the input distribution to compute the output distribution. One form which such a "computation" can take is to employ a set of inputs having weights or frequencies which approximate the input distribution and actually to "physically compute" the response distribution by performing an "aggregated experiment". This approach, in effect, is the one followed in the use of the California driving cycle as a basis for emissions tests. An alternative approach might be to isolate various operating "modes" which make up the schedule and to determine the corresponding times in mode. If the emission rates for each mode were determined by test, the emissions for the entire cycle could be approximated by weighting each mode according to its corresponding time in mode. The advantage of such an approach is that it would not be constrained to a single driving schedule; rather, one could compute the effect of vehicle emissions on the atmospheric pollutant burden according to a variety of assumptions about "typical" driving experience.

In summary, two general approaches to aggregated assessment of automotive emissions have been delineated, each implying a different approach to the layout of emissions experiments associated with development of the stratified-charge engine. In one, we endeavor to put together a "representative" set of input conditions and determine the corresponding output set of emissions without attempting to define cause-effect relations on a one-to-one basis. In this approach, the "transfer function" of the "black box" in Figure 1 is ignored on the basis that it is unnecessary to know the effect of specific parameters on emissions so long as their effects are aggregated in accordance with prevailing or realistic aggregation of causes. In the other approach, we determine the input-output relationships. Then, given any input spectrum we can estimate the output spectrum. For some aspects of the development program, the first may be desirable or necessary; for other aspects of the program the second approach may be preferred. The appropriate tradeoff between the two approaches will be borne in mind in subsequent sections of this discussion as they pertain to evolving a plan for engine emissions assessment. A general guideline is that relatively small effects can afford to be statistically aggregated and the factors which induced them relegated to the limbo of "unassignable causes." Relatively large effects, on the other hand, should be traced to their causes, in the interest of both better repeatability of test results and engineering exploitation of the cause-effect relation to minimize emissions.

2. RANDOM FACTORS AFFECTING EMISSIONS

Factors which can appropriately be treated as random contributions in a statistical model often occur as "nested" or "hierarchical" components. For example, in describing an experiment aimed at determining calcium concentration in turnip greens, Snedecor examined replicate chemical analysis within a single leaf, variation among leaves of the same plant, and variation among plants. In the same way that determinations are nested in leaves and leaves are nested in plants, we can consider replicate emission tests conducted by the same operator, differences among operators in the same laboratory, and differences between laboratories. At each level of the hierarchy it is possible for errors to be introduced, so that a general model might be

$$x_{ijk} = \mu + \alpha_i + b_{ij} + \epsilon_{ijk}$$

where α_i = fixed effect of i^{th} laboratory

$$b_{ij} \sim N(0, \sigma_B^2)$$

$$\epsilon \sim N(0, \sigma^2)$$

It is postulated that ϵ_{ijk}, the test-to-test variability within a single operator and laboratory is a random variable and that b_{ij}, the operator component, can be considered a random variable. On the other hand, each laboratory might bring a particular bias to the determination (as a result of instrumentation differences or other causes), so that the laboratory contribution α_i can be regarded as a fixed effect. This model is accordingly a "mixed model" in the sense that some components of the model are fixed while others are regarded as outcomes of random variables.

How can variation attributable to laboratory-to-laboratory and operator-to-operator sources be controlled? Laboratories are often brought into agreement by a system of interlaboratory comparisons, and operator differences can be "averaged out" by replicating the test a number of times, each time with a different operator. Obviously these techniques are expensive and time consuming, and other, more economical approaches are desired.

An apparently straightforward approach to this problem is by the use of a control sample. This control might take the form of a conventional or modified engine the emissions of which have been well established. In any test series, this "standard engine" could be included as a reference point. If a particular operator or laboratory tends to "run high" or "run low" on this standard, test results can be adjusted accordingly.

Though the use of a standard or control engine appears attractive, the approach is not without problems. Engines, even conventional ones, differ among themselves and will operate differently from time to time for reasons other than those associated with operator response and laboratory-dependent causes. For that reason, the concept of a standard or control engine provides something less than a "stable platform" on which to base results. An approach which should be investigated is the use of an "internal standard" slaved to the system under test, as shown in Figure 2.

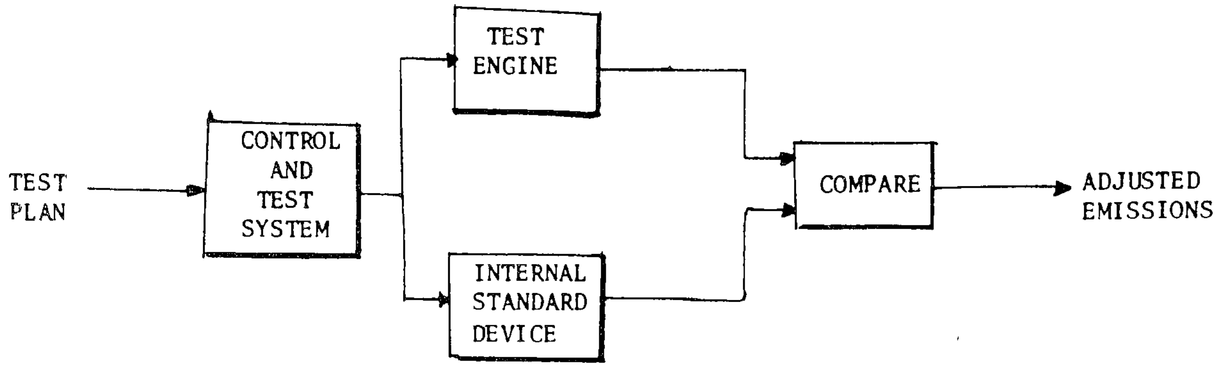


FIGURE 2
STANDARDIZATION OF EMISSION TESTS

The internal standard device could be an engine subjected to the same controls as the test engine, but more appropriately it might be an 'engine simulator' designed to 'emit' pollutants such as CO and hydrocarbons under known and precise control in accordance with operator demand. These effluents could be precisely metered, so that an engine simulator in one laboratory could be expected to perform very closely like one in another laboratory. Differences from one operator to another would be registered as differences in output of the standard device and, by comparison of the test engine output with the simulator output, an emission measurement adjusted for operator idiosyncrasies could be achieved. It is even feasible that the simulator could be entirely mathematical a device which integrates operator demands in such a way that the standard output could be computed rather than being physically generated.

3. NON-RANDOM FACTORS AFFECTING EMISSIONS

Uncontrollable variables which inadvertently enter an emissions test and make exact repeatability impossible are clearly to be regarded as random effects. Operator variables, it has been seen, are of such a nature that they might be treated as either fixed or random effects, depending on whether it is more practical to isolate operator idiosyncrasies and adjust for them or to average them out by statistical aggregation via replication. Laboratory differences, on the other hand, are often rather clearly of a fixed nature, constituting a bias in results. Similarly, it has been seen that influences which might seem to be clearly of a fixed variety, such as the effect of inertial weight, can be couched in a statistical framework as an induced statistical distribution. Thus it is apparent that any discussion of non-random factors affecting emissions must be interpreted in a non-absolute way. Though there are a number of variables which "usually" are to be regarded deterministically, practically all of them are capable of statistical aggregation under suitably defined ground rules.

Perhaps the most clear-cut case of a fixed or non-random factor is the set of design elements which differentiates the Ford from the Texaco stratified-charge engine. This observation is not meant to imply a large engineering difference in the two concepts but rather to emphasize that we are dealing with a dichotomy and that if we are to consider the two concepts statistically they constitute two elements which exhaust the entire population. It is clear, therefore, that a variable labeled "make of engine" and having two "levels", Ford and Texaco, provides a starting point for an experiment-design "tree".

Next, how shall the engines be tested? In a military or civilian vehicle, with standard or automatic transmissions, or by means of an engine dynamometer? What catalyst shall be used, what fuel(s), and what basis for EGR? These and many other variables have been recognized by EPA as candidates for parametric variation to determine their corresponding "fixed effects" on emissions. These major variables are listed in Table II.

TABLE II

MAJOR VARIABLES

Supplier:	Ford	Texaco
Transmission/ Drive:	Stick-4/Auto-2	Stick-4
Catalyst?	Yes or No	Yes or No
Weight:	3000 or 2000	3000 or 2000
EGR:	None/Design/High	None/Design/High
Fuel:	Gas	Gas/Diesel
Aspiration:	Nat.	Nat./Turbo

Not included are such "engine personality" items as age (mileage), state of adjustment or tuning, and individual manufacturing variations. It is such items as these which take on a "borderline" role as far as whether they are to be regarded as isolatable, fixed effects or must be statistically aggregated as random effects. These two classes of variables will be addressed separately in the following discussion.

3.1 Parametric Variables

First, let us consider that set of variables in Table II, which can be considered as variables which we wish to explore parametrically to determine their effect on emissions. These variables will be referred to as "treatment" variables, and, if a level of each variable is specified, the resulting set of constraints will be called a "treatment". For example, one treatment might be the Ford engine with automatic drive, without catalyst, weight 3000, high EGR, gasoline fuel and natural aspiration.

Consider an experiment in n treatment variables $x_1, x_2, \ldots, x_n$. Let x_1 assume k_1 levels, x_2 assume k_2 levels, ..., and x_n assume k_n levels. Denote these sets of levels as:

$$S_1 \quad \{x_{11}, x_{12}, x_{13}, \ldots, x_{1k_1}\}$$

$$S_2 = \{x_{21}, x_{22}, x_{23}, \ldots, x_{2k_2}\}$$

$$S_m = \{x_{m1}, x_{m2}, x_{m3}, \ldots, x_{mk_m}\}$$

Then the Cartesian product $S_1 \times S_2 \times \ldots \times S_n$ constitutes a factorial experiment consisting of $K = k_1 \; k_2 \ldots k_n$ points in treatment space. That is, a factorial experiment is made up of all combinations of levels in which one level is taken from each factor. It can be considered as a starting point for attempts to develop an "economized" version of an experimental program to adequately explore the treatment space. The desirability of such economization becomes evident when it is noted that

2 suppliers x 2 transmissions x 2 catalysts
x 2 weights x 3 EGR options = 24 treatments

and that each treatment would require replication in order to assess random effects or errors.

In order to appreciate the logic by which the above experiment can be abbreviated, it is instructive to consider the matter of interaction between variables. Suppose, for example, that it is possible to express emissions in terms of the equation

$$y = b_0 f_0(x_1, x_2, \ldots, x_n) + b_1 f_1(x_1, x_2, \ldots, x_n) + \cdots + b_k f_k(x_1, x_2, \ldots, x_n)$$

The function f_0, f_1, $\ldots f_k$ can be of any form but they are often considered as polynomials of increasing degree in the x_i. For example, consider a 2 x 2 factorial experiment (two factors each at two levels), often called a 2^2 factorial, in the two variables x_1 and x_2. Then it is customary to take the model as

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2$$

where the subscripts of the regression coefficients are assigned to conform to the variables involved.

The <u>main effect</u> of a factor is defined to be the average effect of that factor on the yield or response variable, where the average is taken over all levels of the other variables. In quantitative experiments, the main effect of x_i corresponds to $b_i = \frac{\partial y}{\partial x_i}$, in the event that there is only one degree of freedom (2 levels) for the variable x_i. If there are <u>three</u> levels for the factor, we can compute an <u>average quadratic main effect</u> as $b_{ii} = \partial^2 y / \partial x_i^2$ and similarly for higher order terms.

A two-factor <u>interaction</u> (or first-order interaction) tells how the effect of one factor depends on the levels of one or more other factors. For example, in

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_{12} x_1 x_2$$

we have

$$\frac{\partial y}{\partial x_1} = b_1 + b_{12} x_2$$

which clearly shows that the effect produced by x_1 on y may be modified by the value assumed by x_2. Thus,

$$\frac{\partial}{\partial x_2}\left(\frac{\partial y}{\partial x_1}\right) = b_{12}$$

measures the interaction between the treatment variables x_1 and x_2. If there are more than two factors in the experiment, the (x_1, x_2)-interaction is the average value of $\partial^2 y / \partial x_1 \partial x_2$, where the average is taken over all levels of the other variables.

For example, in a 2^3 factorial experiment, with the model

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + b_{12} x_1 x_2 + b_{13} x_1 x_3 + b_{23} x_2 x_3 + b_{123} x_1 x_2 x_3$$

we have

$$\frac{\partial y}{\partial x_1} = b_1 + b_{12} x_2 + b_{13} x_3 + b_{123} x_2 x_3$$

and

$$\frac{\partial^2 y}{\partial x_1 \partial x_2} = b_{12} + b_{123} x_3$$

Then the first-order interaction of x_1 and x_2 is the value obtained when $\partial^2 y / \partial x_1 \partial x_2$ is evaluated at all levels of x_3 and averaged. Upon further differentiation, one obtains

$$\frac{\partial}{\partial x_3}\left(\frac{\partial^2 y}{\partial x_1 \partial x_2}\right) = b_{123}$$

which, in this case, measures the three-factor interaction (second-order interaction) of x_1, x_2, and x_3. It shows the extent to which the first order interaction of x_1 and x_2 depends on the level of x_3. If there are more than two levels in the treatment variables, we can have, in addition to the linear x linear interaction above, such interactions as linear x quadratic, and quadratic x quadratic (and so on).

A factorial experiment is designed to estimate all interactions, up to the maximum degree possible within the scope of the design. It may often be evident, however, from engineering analysis and experience that certain interactions are unlikely to be of any significant magnitude. In such cases, the experiment can be redesigned on the assumption that the interactions in question are non-existent. This approach allows simplification of the experiment and the elimination of certain treatments, the result being a more economical design. A typical example is provided by the class of so-called "fractional factorial" designs, in which only a fraction-- say 1/2 or 1/4-- of all the treatments in the full factorial are employed. An alternative approach is provided by designs of the type developed by G.E.P. Box and his colleagues, called response-surface designs. These designs are developed so that interactions of little importance are ignored, but those likely to be important are not biased by the simplification of the design.

It is proposed to employ these principles, although not necessarily formally, in the formulation of a testing program for stratified charge engines. For example, consider the transmission/drive variable and such variables as catalyst and EGR. If it can be postulated that the transmission/drive variable does not interact with the catalyst or EGR variables, then it would be unnecessary to run through the gamut of EGR and catalyst in both automatic and non-automatic transmission modes. Rather, it might suffice to run through the gamut of these variables in either automatic or non-automatic mode and obtain only a single or "bench-mark" point in the other mode. Note that the assumed absence of interaction implies parallel responses in the two transmission modes but allows for the possibility that the response in one mode might be displaced from the response in the other mode, as shown in Figure 3.

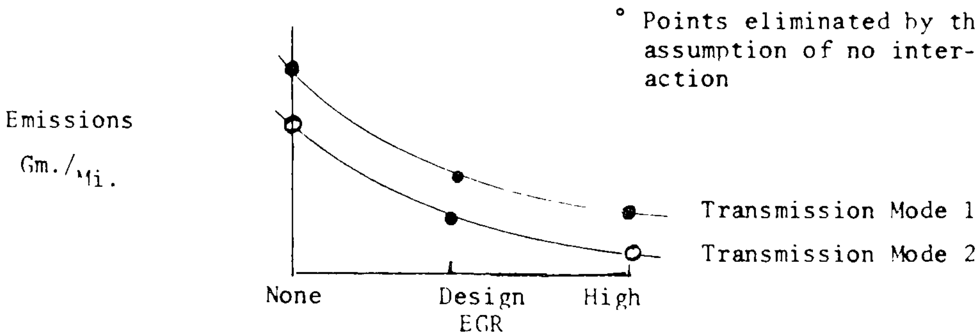


FIGURE 3
ILLUSTRATION OF LACK OF INTERACTION

Thus two of the required six test points could be eliminated without seriously jeopardizing the validity of conclusions which might be drawn from the testing program. If one pushes this approach even further, it will be apparent that many variations in engine operating variables could be explored by tests on an engine dynamometer, especially if it could be shown that the vehicle configuration environment tends to introduce only differences of the "bench-mark" variety.

3.2 Variables Amenable to Statistical Aggregation

Those variables to be considered for statistical aggregation are those which are associated with engine "personality." These include, for example, minor deviations in dimensional tolerances, day-to-day vagaries of engine performance, and the like. Admittedly, the effects of many of these variables could be studied by parametric variation. For example, one could deliberately introduce dimensional variations and observe their consequences on emissions. It is considered more practical, however, to undertake statistical aggregation by testing a large enough sample to allow statistical variation to come into play.

Consider a collection of N engines of a particular class and let these engines be tested repeatedly. It will be assumed that the test laboratory is fixed, that the same operator is used, and that all other testing variables which might introduce variation have been eliminated. Nevertheless, it will be observed that successive tests of a particular engine do not give identical results. It has been said that no one can step in the same river twice. In the same vein, it can be said that no one can test the same engine twice. It has "aged" by some finite increment, if nothing else, and therefore may not perform in exactly the same way as it did originally. As the tests continue, other influences will occur which will also cause some degree of variation in test results--quite apart, please note, from testing errors introduced by instrumentation and operator effects. Some of these influences, such as the aging effect noted above, may be capable of being treated deterministically. Others will need to be treated as random variables.

In previous attempts to separate engine-to-engine variability from test-to-test variability, it has been assumed that repeat tests performed on the same vehicle provided a measure of such test-associated variation as that arising from instrumentation sources, deviations from the driving schedule, etc. It must be recognized, however, that there are within-engine sources of variation which are associated with the engine itself: wear of parts, idiosyncrasies of linkages, and the like. It is also quite conceivable that the day-to-day or time-to-time differences of this nature within a particular vehicle may be more pronounced than differences from one vehicle to another. It is for this reason that repeated testing of a given vehicle over a considerable period of time and with the vehicle in a range of normally encountered operating conditions is in order. A more realistic variance model is therefore

$$x_{ijk} = \mu + a_k + b_{jk} + \epsilon_{ijk}$$

where a_k is a contribution associated with the k^{th} vehicle, b_{jk} is a <u>vehicle-associated</u> contribution of the j^{th} test on that vehicle, and ϵ_{ijk} is a <u>test-associated</u> error contribution. The question naturally arises as to whether the error ϵ_{ijk} can be dissociated from the vehicle-associated contribution b_{jk} or whether these two sources of variability are confounded in such a way that they can not be separated. We believe, however, that if replicate tests are run on the same vehicle at as nearly the same time as possible, this procedure will tend to eliminate time-to-time variations in the engine itself. On the other hand, if the engine is tested over a period of time and conditions, again each time in replicate, the two sources of variability can be, in effect, separated.

In assessing the b_{jk}, one must consider whether these contributions to the variability of x_{ijk} are random or fixed effects. In some cases the effects may be considered to be of the fixed variety. For example, in the case of engine age or accumulated mileage, it is possible that a time trend in emissions might be observed, especially if the vehicle is operating with a catalyst subject to deterioration with time. Such trends could be extracted by means of regression analysis and the remaining or residual variance treated as arising from random causes.

4. TESTS ON AVAILABLE HARDWARE

The discussion of the previous sections has been regarded as necessary considerations prerequisite to the formulation of a testing program for stratified charge engines. Most important is a careful weighing of what types of testing are appropriate to existing hardware and what types of testing should be reserved for new hardware arising from the development program.

In this connection, it must be acknowledged that with only a few vehicles available, it is virtually impossible to do much in assessing the engine-to-engine variability which arises from manufacturing variations or associated causes. It is, however, quite logical to employ these few vehicles in a program aimed at defining the range of emissions which can occur from variations within the engine itself. For example, a program aimed at mileage accumulation coupled with routine maintenace and adjustment could do much to define the range within which a specific engine might be expected to operate, as far as emissions are concerned. It is also conceivable that parametric variation of major variables could be undertaken to define emission response to these variables. If the latter course is followed, however, it might be well to ask whether that type of study is premature, in view of the fact that design concepts are still in a state of flux. We recommend, therefore, that parametric tests be undertaken with a view toward defining primarily the direction of trends rather than attempting to measure with high quantitative accuracy the actual magnitude of these trends. With this thought in mind, we consider it adequate to assess the effect of such parameters as inertial weight, EGR and catalyst only at the beginning and end of the mileage accumulation program. Intermediate tests could be performed on some single'bench mark treatment', such as 2000 lbs. weight, design EGR, and standard catalyst. This single configuration would make it possible to follow time trends in emissions which might be caused by engine deterioration. Though it is possible that there is interaction between treatment configuration and mileage, an assessment of these interactions would be available from the parametric tests to be performed at the beginning and end of the mileage-accumulation program. The approach would have the advantage that intermediate tests could be expedited, since no variations in test setup would be required. In this way, the time taken away from mileage accumulation would be minimized and yet the loss in parametric information would be minimal. Though intermediate tests might serve to strengthen confidence in the effect of treatment parameters on emissions, a similar strengthening of confidence could be attained by running additional replicates at the end of the mileage-accumulation program.

The testing program outlined in Table III is proposed for evaluating existing hardware, specifically one 4-wheel drive Jeep equipped with Ford Combustion Process (FCP) engine, one 4-wheel drive Jeep equipped with Texaco Combustion Process (TCP) engine, and one 2-wheel drive postal van, equipped with FCP engine and automatic transmission. Bench-mark tests of a standard Jeep equipped with a conventional engine are included as points of reference against which the stratified-charge engines can be compared. It is proposed that the standard Jeep tested in the initial characterization be in substantially new condition and that the Jeep tested in the final characterization be one with substantial (approximately 50,000) mileage accumulation. These two tests need not be performed on the same Jeep, since the intent is only to provide a basis for comparison of the stratified charge engines with "typical" conventional vehicles. It is proposed that each test be conducted in duplicate so as to provide a basis for assessment of errors. Preferably the replicate tests should be performed by different operators drawn randomly from a pool of several operators taken as representative of operators in general, so that operator variability will be reflected in the repeatability assessment. Otherwise, the measured repeatability might be unrealistically optimistic.

The test array of Table III is regarded as being "adaptive" in the sense that it may be modified as the program progresses. For example, it is indicated that mileage accumulation to 50,000 miles is projected with emissions tests scheduled at approximately 10,000 mile intervals. In the

TABLE III

EMISSIONS TESTING PROGRAM FOR STRATIFIED CHARGE ENGINES

	JEEP FCP - JEEP TCP - POSTAL VAN TCP												STANDARD JEEP
	CATALYST						NO CATALYST						
	2000# Inertial Weight			3000# Inertial Weight			2000# Inertial Weight			3000# Inertial Weight			
	No EGR	Design EGR	High EGR	No EGR	Design EGR	High EGR	No EGR	Design EGR	High EGR	No EGR	Design EGR	High EGR	
Initial	H	C	H	H	C	H	H	H	H	C	H	H	C
10,000 Mi.					C								
20,000 Mi.					C								
30,000 Mi.		C			C					C			
40,000 Mi.					C								
50,000 Mi.					C								
Final	H	C	H	H	C	H	H	H	H	C	H	H	C
Rejuvenated		C			C								

H = Hot start only

C = Complete cold-start test

the event that changes either in performance or emissions are observed to occur rapidly with increasing mileage, it may be desirable to alter either the extent of the mileage accumulation or the frequency of testing. In short, it is contemplated that as information accumulates from the testing program, this information will be used as feedback to the experiment design.

The test array is not structured according to any particular maintenance schedule. In view of the lack of experience with stratified-charge engines, it is believed that to do so would be unrealistic. Rather, it is contemplated that whatever maintenance is required would be conducted as necessary and that any effects introduced by such maintenance would be treated either as fixed or random effects. For example, if major or at least definitive repairs were necessary, such as the changing of spark plugs or replenishment of catalyst, these would be noted in the testing log and possibly adjusted for as fixed effects in the data analysis. On the other hand, such routine maintenance as tuning, adjustment of linkages, etc., being less definitive, would be aggregated as additional statistical variation or error in the test results.

A word of explanation is in order concerning the tests scheduled as "Rejuvenated." The intent here is to see how much of the observed degradation in emission characteristics is of a reversible nature. It is proposed to replace catalyst and spark plugs, retune the engines, and perform any other simple maintenance which would "give the engine every benefit of the doubt" as far as emissions are concerned. In short, one would endeavor to see to what extent the engine could be "brought back" to its initial emissions by means which could be readily and simply implemented.

In proposing the s edule of complete (C) and hot-start-only (H) emissions tests, we were governed by considerations of economy and the experiment-design principles previously discussed. In effect, the two extensive series of tests designated as Initial and Final can be considered as two levels of a factorial experiment in which the variable in question is mileage and the levels are nominally 0 and 50,000. Since these two levels represent extremes, it is considered that the effects of other variables, such as inertial weight, EGR and catalyst, may be different before and after the mileage accumulation. In short, it is considered that these variables may interact with the mileage-accumulation variable. Also, though it is realized that intermediate mileages may show similar effects, it is considered adequate to monitor the development of these effects by a single "treatment", 3000# weight, design EGR with catalyst. This treatment is regarded as a design baseline and, as such, is appropriately studied in greater depth than off-design variations. Moreover, this single baseline case can serve to monitor changes in mileage and serve as an indicator or "sensor" the output of which can be used as feedback to the testing strategy. If the results of the design baseline tests indicate rapid degradation, this information can be used to implement an earlier final assessment or a more complete intermediate set of parametric tests.

Similar reasoning applies to the selection of treatments for complete tests and for hot-start-only tests. It is considered that the relative effects of changes in such variables as catalyst and EGR can be observed through hot-start only tests. On an absolute basis, however, it is realized that the emissions observed on hot-start-only tests will be lower than if the cold-start emissions are included. Realism, as far as the general level of emissions is concerned, can be recovered by running one treatment, the design baseline case, under the complete test. By eliminating the other cold-start cases, one effects considerable saving in time and costs. Abbreviation of the testing program for the no-catalyst cases to a single treatment for complete test is justified on the basis that the presumably worst-emissions case is selected. This single case provides a "bench-mark" for the no-catalyst cases and presumably would be most like the standard jeep. This case would thus provide, together with the conventional engine, an assessment of what improvement in emissions might accrue from the stratified-charge concept alone. This comparison is supplemented by parametric variations conducted by hot-start-only tests within the no-catalyst subset of treatments.

APPENDIX I NEW HARDWARE REQUIREMENTS FOR ASSESSING EXHAUST EMISSIONS PERFORMANCE OF THE STRATIFIED CHARGE ENGINE A STATISTICAL CONSIDERATION

H. T. McAdams

One of the most important considerations in any hardware development program is the question of how many units must be built and tested in order to have "reasonable assurance" that the developmental item will perform satisfactorily. The question is especially serious if fabrication costs and time schedules limit the availability of units for testing purposes. In such instances, it is necessary to draw conclusions from a limited amount of data but yet have some measure of the validity of those conclusions.

The nature of the question is at least twofold. First, one must define in quantitative terms what is meant by reasonable assurance. Second, there must be some inferential process by which the results obtained from a limited number of units can be projected to what might be expected when the experimental hardware is produced in larger quantities.

1. CONFIDENCE INTERVALS

The above question is only a special case of one of the most persistent problems in statistical inference: how much data is enough? Because of the statistical nature of the question, its answer is often expressed in terms of a confidence interval. Properly interpreted, the concept of confidence interval is a useful one in answering the question of "How many observations are needed to establish the mean value of a random variable to within certain prescribed limits and to a certain level of confidence?"

However, since the implications of a confidence interval can be misinterpreted, it is in order to discuss the concept briefly in the context of stratified charge engine development.

Let us imagine that a very large supply of engines is available. Suppose that N of these engines are selected at random, where N is a small number relative to the total number available, and that each engine in this small sample is tested once for its emissions per mile. From the N tests, the mean or average value can be computed, as can also the standard deviation of the individual measurements about this mean. Since the N engines measured are only a "sample" of the entire population of engines, a second set of N engines will not necessarily yield the same sample mean or sample standard deviation. Moreover, neither of the sample means are likely to coincide with the population mean, which could be determined by measuring emissions for "all" engines in the population. Similarly, the sample standard deviations would be unlikely to coincide with the "true" or population standard deviation similarly obtained by exhaustively measuring all engines in the population. The concept of "confidence interval" represents an attempt to invest a sample result with some measure of "nearness" to the population result. For large values of N, for example, it can be shown that the sample mean will tend to be "closer" to the population mean than would be the case for small values of N. By the reverse of this reasoning process, one can postulate how large N should be in order to achieve a certain degree of "nearness" to the true or population mean. It is evident, however, that the question of "How many tests are needed?" must be preceded by a question which asks, in essence, "How close to the truth is close enough?"

The mechanics of computing a confidence interval is straight-forward. For example, suppose that 10 engines are tested and that for each engine the hydrocarbon emissions per mile is determined. Their mean is

$$\bar{x} = \frac{1}{10}(x_1 + x_2 + \cdots + x_{10})$$

and their sample standard deviation* is

$$s = \sqrt{\frac{\sum_{i=1}^{10} (x_i - \bar{x})^2}{10}}$$

As a random variable, $\bar{x}$ is an unbiased estimator of μ, the population mean. Similarly,

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{N} (x_i - \bar{x})^2}{N - 1}$$

is an unbiased estimator of the population variance.

Inasmuch as $\bar{x}$ is a random variable, it has its own variance; the magnitude of this variance is inversely proportional to N, the sample size. Thus an unbiased estimator for the variance of $\bar{x}$ is

$$\frac{\hat{\sigma}^2}{N} = \frac{\sum_{i-1}^{N} (x_i - \bar{x})^2}{N(N-1)}$$

and the square root of this quantity is the standard error of the sample mean. By making certain assumptions regarding the statistical distribution of $\bar{x}$ one can compute the required confidence interval.

---

*The sample variance, defined as $s^2 = 1/N \sum_{i=1}^{N} (x_i - \bar{x})^2$ does not provide an unbiased estimate of σ^2, the population standard deviation. The bias is eliminated by using N-1 in place of N.

Confidence intervals rely on the assumption that the sample mean, as a random variable, obeys a normal or Gaussian distribution with some population mean μ and variance σ^2/N, where σ^2 is the population variance of X and N is the sample size.* If repeated samples are taken from the population, and if σ^2 is known, approximately 68% of the resulting values of $\bar{x}$ will fall within an interval bounded on the low side by $\mu - \sigma$ and on the high side by $\mu + \sigma$. Similarly, approximately 95% of the values of $\bar{x}$ will lie in the interval $\mu - 2\sigma$ to $\mu + 2\sigma$ and other percentages of the cases can be bracketed by appropriate choice of k in the quantities $\mu - k\sigma$ and $\mu + k\sigma$ which bound the interval. Suppose, however, that only the result from a particular sample is available. The sample quantities $\bar{x}$ and $\hat{\sigma}$ provide estimates of μ and σ, and it is presumed that these values can be used to construct a confidence interval

$$\bar{x} - k\hat{\sigma} \leq \mu \leq \bar{x} + k\hat{\sigma}$$

and that by proper choice of k, the width of the interval can be adjusted in such a way that there is 95% confidence that the population mean μ will be contained in the interval. For example, if $\bar{x}$ = 0.32 grams per mile and $\hat{\sigma}$ - 0.06 grams per mile, a 95% confidence interval for μ is provided by

$$0.32 - k\left(\frac{0.06}{\sqrt{N}}\right) \leq \mu \leq 0.32 + k\left(\frac{0.06}{\sqrt{N}}\right)$$

For large values of N, the value of k approaches 1.96, but for small values of N, the value may be considerably larger. For example, if N 9, k - 2.306, and the interval becomes approximately

$$0.27 \leq \mu \leq 0.37$$

---

* If N is only moderately large, the validity of this assumption is assured by the Central Limit Theorem of mathematical statistics. In statistical quality control for example, it is found that N may be as small as 4 or 5.

A word of caution is in order, however. The population mean is a constant, not a random variable. In reality, therefore, μ either is or is not contained in the above interval. It is the interval, as computed from the sample data, which statistically varies; both its midpoint and its width depends on the particular values obtained in the sample. If repeated samples of 9 engines were tested and a confidence interval computed for each of these samples, approximately 95% of the intervals so constructed would contain the population mean μ . It is in this sense that the confidence interval should be interpreted. Thus, in the above example, if one asserts that μ falls in the interval 0.27 to 0.37, such an assertion is "more likely" to be true than false; in the sense indicated, the assertion has probability 0.95 of being a true statement.

It will be noted that the width of the confidence interval depends on several factors: the level of confidence, the standard deviation of the measurements, and N, the number of engines tested. The level of confidence is often set arbitrarily; however, its choice can be made rationally by examining the consequences of exceeding the interval bounds. The standard deviation is determined in part by the variation which exists from one engine to another and in part by the inability to repeat successive emission measurements exactly because of experimental errors. The sample size, N, is determined by availability of engines and, more importantly perhaps, by procurement costs.

2. VARIANCE COMPONENTS

As pointed out above, there are at least two sources of variance which must be reckoned with in the assessment of automotive emissions:

a) variance induced by differences among engines

b) variance induced by testing error.

Thus a model for emissions might be*:

$$y_{ij} = \mu + \alpha_i + \beta_{j(i)}$$

where μ contribution common to all measurements

α_i contribution peculiar to the i^{th} engine

$\beta_{j(i)}$ = contribution peculiar to the j^{th} test of the i^{th} engine

In any single test of an engine, the variance σ^2 can be expressed as

$$\sigma^2 \quad \sigma_e^2 + \sigma_t^2$$

where σ_e^2 component of variance attributable to engine-to-engine variability

σ_t^2 component of variance attributable to test-to-test variability.

* A more general model, however, would further delineate the identifiable sources of testing error, such as operator error, and the sources of differences among engines, such as production tolerances, adjustments, etc.

The distinction is an important one because it has direct bearing on the strategy of any testing program aimed at predicting the performance of engines as a population. If the engine component σ_e^2 is a small fraction of σ^2 there is little need to sample extensively the engine population. Consequently, in this case, statistical needs could be satisfied by testing only a few randomly selected engines. It would be necessary to perform a large number of repeat, or replicate, tests on each engine, however, because the testing variability would constitute the major fraction of σ^2. Suppose, on the other hand, that σ_e^2 is a large fraction of σ^2. Then a relatively large sample of engines would be necessary, but a few tests on each would suffice because test-to-test variation would make only a small contribution to the total variance σ^2. It must also be noted that if σ_e^2 is large, no amount of replicate testing on a few engines will strengthen our confidence substantially because it provides no mechanism for averaging differences among engines, the main source of variability.

3. SAMPLE SIZE RELATIONSHIPS

In many testing situations, the magnitude of the measurement errors tends to be proportional to the mean value of the quantity being measured. Experience with both automotive and aircraft engine emissions suggests that this relationship applies, at least to good approximation, in the case of gaseous emissions measurements. It is convenient and pertinent, therefore, to express the standard deviation of the measurements as a fraction of their mean. This ratio is often referred to as the coefficient of variation, or relative standard deviation.

Let $\bar{x}$ denote the sample mean for a sample of N engines selected at random, each engine being tested once. Suppose that the population mean is μ and that the population standard deviation is

$$\sigma = p\mu$$

where p is the coefficient of variation. The standard error of the random variable $\bar{x}$ is accordingly

$$\frac{\sigma}{\sqrt{N}} = p\mu/\sqrt{N}$$

If it is presumed that σ is known, then a 95% confidence interval for μ is given by

$$\bar{x} - 1.96\ p\mu/\sqrt{N} \leq \mu \leq \bar{x} + 1.96\ p\mu/\sqrt{N}$$

This result appears circuitous in that μ would need to be known in order to establish the confidence interval, but if it were known there would be no need for a confidence interval. Suppose, however, that one considers only the width of the interval, which is

$$\bar{x} + 1.96\ p\mu/\sqrt{N} - (\bar{x} - 1.96\ p\mu/\sqrt{N}) = 3.92\ p\mu/\sqrt{N}$$

Then the width of the interval, expressed in terms of μ, is just

$$3.92\ p\mu/\sqrt{N} \div \mu = \frac{3.92\ p}{\sqrt{N}}$$

or the interval half-width is 1.96 $p/\sqrt{N}$.

This simple result can afford broad guidance in postulating the number of tests required to establish a certain degree of confidence in emissions measurements. It should be noted that a similar argument can be developed for other levels of confidence, say 99%. The 95% confidence level was selected, however, as the basis for further analysis. Confidence intervals based on the 95% level of probability provide a high degree of assurance that a particular interval will contain the population mean. Beyond 95%, the width of the interval (in terms of multiples of the standard deviation) increases rapidly for each additional 1% increase in confidence.

In Table 1 is tabulated the quantity 1.96 $p/\sqrt{N}$ for various values of p and N. The parameter p is varied in 15% increments from 15% to 75%; the parameter N is varied from 1 to 100 in intervals which are perfect squares. According to the table, for data with precision such that the coefficient of variation is 15%, approximately 9 tests would be required to reduce the half-width of the 95% confidence interval to 10% of the population mean, μ. Similarly, if the coefficient of variation is 30%, 36 tests would be required for the same purpose. If the coefficient of variation is as large as 45%, 81 tests would be necessary to determine the mean value to within 10% of its magnitude.

FACTORS FOR CONFIDENCE INTERVALS

95% Confidence Bounds

Sample Size, N	Coefficient of variation, p (%) 15	30	45	60	75
1	29.4	58.8	88.2	117.6	147.0
4	14.7	29.4	44.1	58.8	73.5
9	9.8	19.6	29.4	39.2	49.0
16	7.4	14.7	22.0	29.4	36.8
25	5.9	11.8	17.6	23.5	29.4
36	4.9	9.8	14.7	19.6	24.5
49	4.2	7.7	12.6	16.8	21.0
64	3.7	7.4	11.0	14.7	18.4
81	3.3	6.5	9.8	13.1	16.3
100	2.9	5.9	8.8	11.8	14.[illegible]

TABLE 1

4. IMPLICATIONS FOR COMPLIANCE WITH EMISSION STANDARDS

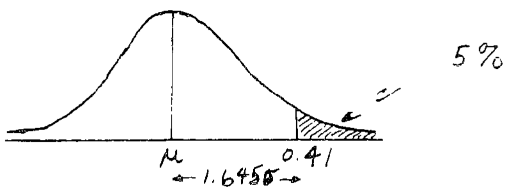
The actual engineering requirements imposed on an engine in order to meet emissions standards is determined, in part, by the manner in which compliance with the standard is to be interpreted. To make this point clear, we consider two cases representing different but reasonable interpretations of the standard.

4.1 Case 1. Compliance Based on a Single Test

Suppose that compliance is based on a single test of a single vehicle. Now obviously if the vehicle is to be passed on the basis of a single test, the expected value or mean for that vehicle must be lower than the nominal standard. One can calculate how much lower by assuming some confidence level say 95% at which to work.

Suppose, for a particular engine, the expected value for all tests is μ and the standard deviation is σ. Then, for that engine, if it is to pass with 95% confidence on the basis of a single test, we must have

$$\mu + 1.645\ \sigma \leq 0.41$$



If σ p% of the mean,

$$\mu + 1.645\ p\mu \leq 0.41$$

and the upper bound is given by

$$\mu(1 + 1.645\,p) = 0.41$$

or, $$\mu = \frac{0.41}{1 + 1.645\,p}$$

where p is the coefficient of variation.

One can assume various values for p and see the consequences. (See Table 2).

4.2 Compliance Based on Expected Value

CASE 2

On the other hand, suppose that the expected value for the engine is not to exceed the standard value 0.41. Then the tolerance for an individual test should be augmented by 1.645 σ. An estimate of σ would be required, but, once known would prescribe that

$$\text{Individual test} \leq 0.41 + 1.645\,\sigma$$

Again, suppose that σ $p\mu$

Then,

$$0.41 + 1.645\,\sigma = 0.41 + 1.645\,p\mu$$

But we want μ not to exceed 0.41. Then,

$$0.41 + 1.645\,p\mu \quad 0.41 + 1.645p\,(0.41)$$

$$- \quad 0.41 + 0.67445\,p$$

where p = coefficient of variation.

This would lead to Table 3.

TABLE 2

VALUES FOR μ FOR INDIVIDUAL ENGINE REQUIRED TO

PASS STANDARD 0.41 @ 95% CONFIDENCE

Coefficient of Variation p	Expected Value For Engine μ
0.00	0.41
0.05	0.38
0.10	0.35
0.15	0.33
0.30	0.28
0.45	0.24
0.60	0.21
0.75	0.18

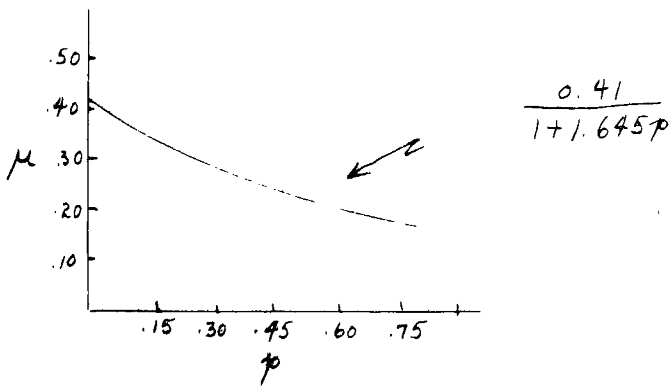
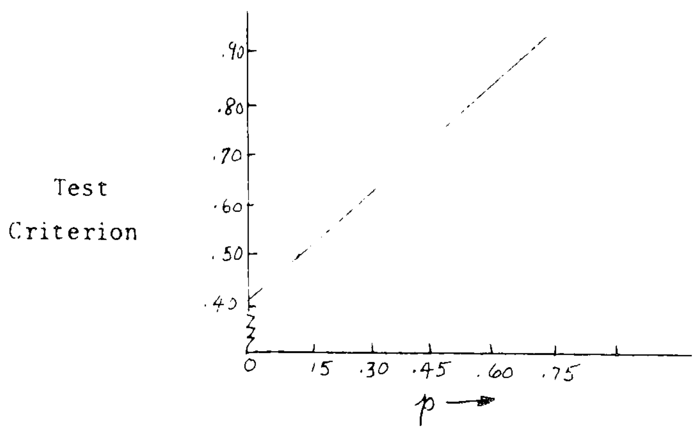


TABLE 3

ALLOWABLE TEST CRITERION TO ASSURE THAT $\mu \leq 0.41$ FOR VARIOUS VALUES OF p

Coefficient of Variation p	Maximum Allowable Single Test Criterion
0.00	0.41
0.05	0.44
0.10	0.48
0.15	0.51
0.30	0.61
0.45	0.71
0.60	0.82
0.75	0.92



★ U.S. Government Printing Office: 1973--747-787/322 Region No. 4

It should be noted that the desired emissions level (say 0.41 grams per mile) and the basis on which an engine should pass are not and need not be the same. An acceptance plan must reflect the desired statistical risks and could be single sample, multiple sample, etc. For example, if a single test does not pass, a retest could be allowed under conditions adjusted to give a prescribed statistical risk. The procedural basis for various sampling plans is well established and could provide a sound statistical basis for compliance testing.

5. IMPLICATIONS FOR TECHNICAL FEASIBILITY OF A STRATIFIED CHARGE LOW EMISSIONS ENGINE

It is seen that, whatever the basis of compliance with standards, the expected values of the emissions for a particular engine must be known if one is to assess the technical feasibility of producing a low emissions stratified charge engine. The term feasibility is interpreted to mean that it is not sufficient to produce a single engine capable of meeting emissions standards but that it should be possible to produce such engines in quantity without incurring excessive failure rates.

One way to approach this problem is to assume that the testing component of variance, σ_t^2, can be substantially eliminated from consideration by performing a sufficient number of replicate tests on each engine. This assumption presumes, therefore, that it will be possible to determine the expected value of emission for each pollutant for each engine tested. One must be concerned, therefore, only with engine-to-engine variability. The problem is to determine what the expected emissions values must be when expectation applies to the entire engine population, if it is desired to reduce rejection rates to any prescribed value.

Suppose, for example, that we do not want more than 5% of all engines to exceed 0.41 gms/mile. Then, if σ_e is the standard deviation of engine expected values about the population expected value, we must have

$$\mu + 1.645\ \sigma_e = 0.41$$

or, if σ_e $p\mu$

$$\mu = \frac{0.41}{1 + 1.645\ \sigma}$$

and Table 2 applies (but this time to expected values for engines, not to test values for a particular engine). Thus if p 0.15, the population mean for all engines must be $\leq$ 0.33 to assure that expected values for <u>individual engines</u> will not exceed 0.41.

<u>NOW</u>, how many engines must be tested to "predict" that the population mean will not exceed a certain value -- say 0.33.

If we are to work at a value of N to insure a confidence interval $\mu \pm 10\% \mu$, we must decrease 0.33 by 10% to approximately 0.30. The better we know the population mean, i.e. the greater N, the less demanding will the requirement on the estimated mean be. It should further be noted that the demands placed on engineering also strongly depends on σ_e or more specifically. the coefficient of variation p obtained when σ_e is expressed as a fraction of the population mean for the emission measurement under consideration. If p is taken as 15%, and if it is desired to hold to a confidence interval $\mu \pm 10\% \mu$, then the number of engines to be tested is approximately N = 9. If N 4, the half-width of the interval goes to approximately 15%, and we would have to depress the 0.33 value not to 0.30 but to

$$0.33 \quad (0.15)\ (0.33) = 0.28.$$

What constitutes a good estimate of coefficient of variation is problematical, since there is no direct statistical basis for assessment. There is evidence to show that testing variation is such that σ_t is approximately 15% of the mean for a particular engine. Also, in tests on aircraft engines, it was observed that engine-to-engine variability often was at least comparable to test-to-test variability. It is on the basis of these very tenuous arguments that the above analysis was made assuming a 15% coefficient of variation.

In conclusion, it appears that if one were to speak in terms of orders of magnitude for the number of engines to be tested, this number should be at least 10. Though it is suspected that this number is conservative, it would at least provide a basis for a revised estimate of testing requirements. If it should turn out that a larger sample appears necessary, additional engines could be procured and tested sequential to the first sample lot. On the other hand, if a larger number of engines were procured unnecessarily, much expense would have been incurred which could not be written off to any substantial gain in information.

ENVIRONMENTAL PROTECTION AGENCY
Technical Publications Branch
Office of Administration
Research Triangle Park, N C 27711

OFFICIAL BUSINESS

AN EQUAL OPPORTUNITY EMPLOYER

POSTAGE AND FEES PAID
ENVIRONMENTAL PROTECTION AGENCY
EPA - 335

SPECIAL FOURTH-CLASS RATE
BOOK

U.S. MAIL

Return this sheet if you do NOT wish to receive this material ☐, or if change of address is needed ☐. (Indicate change, including ZIP code.)