Calspan

Technical Report

STUDY TO DETERMINE EMISSION DETERIORATION FROM EMISSION FACTOR PROGRAM DATA

TASK III FINAL REPORT

Calspan Report No. NA-5542-V-3

Calspan Corporation
Buffalo, New York 14221

Calspan

STUDY TO DETERMINE EMISSION DETERIORATION FROM EMISSION FACTOR PROGRAM DATA

TASK III FINAL REPORT

Calspan Report No. NA-5542-V-3

Prepared for:

ENVIRONMENTAL PROTECTION AGENCY
NATIONAL ENVIRONMENTAL RESEARCH CENTER
CINCINNATI, OHIO 45268

SEPTEMBER 1976
CONTRACT NO. 68-03-0486
FINAL REPORT

Prepared by: A.C. Keller
A.C. Keller, Task Engineer
Systems Evaluation Dept.

Approved by: M.W. Hall
M.W. Hall, Head
Systems Evaluation Dept.

Calspan Corporation
Buffalo, New York 14221

EPA-420-R-76-108

This report was furnished to the Environmental Protection Agency by Calspan Corporation, in fulfillment of Task III of Contract No. 68-03-0486. The opinions, findings, and conclusions expressed are those of the author and not necessarily those of the Environmental Protection Agency.

ABSTRACT

This report on emission deterioration was prepared for the Environmental Protection Agency, Emission Control Technology Division, Ann Arbor, Michigan, under EPA Contract No. 68-03-0486. The purpose of this study is to estimate emission deterioration of in-use vehicles as they accumulate age and mileage. The analysis uses data collected on vehicles tested in the FY 71, FY 72, and FY 73 Emission Factor Programs.

This study assesses the differences in deterioration rates between (a) Model years, over the three fiscal years, (b) Fiscal years, for a specified model year, (c) Fiscal years, over all model years, and (d) Test sites, for a specific model year. This study also includes investigations to determine the best form of the deterioration curve and the most appropriate independent variable, age or mileage. Changes in the SDS Modal Emissions Model as related to mileage are also evaluated.

This report reviews the problems of predicting emissions with available EFP data, including the statistical and engineering significance of the predictions, the danger of extrapolating beyond age and/or mileage ranges, and the shortcomings of the EFP format for predictive purposes.

ACKNOWLEDGEMENTS

Support and encouragement by the Environmental Protection Agency are gratefully acknowledged. Particular thanks go to Lois A. Platte for her guidance and suggestions.

A number of individuals at Calspan Corporation contributed significantly to this effort. These include William R. Fairchild and Jeffrey C. Bernard, who performed the computer analysis, and H.T. McAdams who provided guidance and helpful comments throughout the entire program.

TABLE OF CONTENTS

Section	Title	Page
1	INTRODUCTION	1
	1.1 Background	1
	1.2 Purpose and General Approach	2
2	SUMMARY	3
	2.1 Choice of Deterioration Curve Form and Independent Variable	3
	2.2 Deterioration Analysis Results -- FTP Data	3
	2.3 Deterioration Analysis Results -- SDS Data	5
	2.4 Problems and Recommendations	5
3	DETERIORATION ANALYSIS	7
	3.1 Data	7
	3.2 Data Analysis and Statistical Tests	7
	3.3 Form of the Deterioration Curves	9
	3.4 Independent Variable Investigation	18
4	DETERIORATION ANALYSIS RESULTS	23
	4.1 FTP Data Analysis	23
	4.1.1 Model Year Effects	24
	4.1.2 EFP Fiscal Year Effects for a Given Model Year	24
	4.1.3 EFP Fiscal Year Effects for all Model Years.	28
	4.1.4 Test Site Effects	28
	4.1.5 CID Effects	28
	4.1.6 Mileage Accumulation Rate Effects	34
	4.2 SDS Data Analysis	34
	4.2.1 Observed Bag Values -- Model Year and Test Site Effects	34
	4.2.2 Emission Deterioration as Reflected in the Modal Emission Model Coefficients	42
	4.2.3 Observed Idle Mode Bag Values -- Changes with Mileage	50

TABLE OF CONTENTS (cont.)

Section	Title	Page
5	PROBLEMS AND RECOMMENDATIONS	53
5.1	Using EFP Data to Determine Historical Emission Trends	53
5.2	Using EFP Deterioration Estimates to Predict Overall Emission Levels	55
5.3	Using EFP Deterioration Estimates to Predict Emission Levels of Individual Vehicles	58
5.4	Recommendations for Future Studies	58
REFERENCES		60
APPENDIX I	T-TEST FOR THE EQUALITY OF TWO SLOPES	I-1
II	LINEAR REGRESSION RESULTS FOR FTP DATA	II-1
III	LINEAR REGRESSION RESULTS FOR SDS BAG VALUE DATA AND FOR SELECTED MODAL EMISSION MODEL PARAMETERS	III-1

LIST OF ILLUSTRATIONS

Figure	Title	Page
1	Linear and Quadratic HC vs. Mileage Regressions for 1971 Model Year Data	11
2	Linear and Quadratic CO vs. Mileage Regressions for 1971 Model Year Data	12
3	Linear and Quadratic NOXC vs. Mileage Regressions for 1971 Model Year Data	13
4	Linear and Quadratic HC vs. Age Regressions for 1971 Model Year Data	14
5	Linear and Quadratic CO vs. Age Regressions for 1971 Model Year Data	15
6	Linear and Quadratic NOXC vs. Age Regressions for 1971 Model Year Data	16

LIST OF TABLES

Table	Title	Page
1	Comparison of Linear and Quadratic Models of Emission Deterioration	17
2	Correlation Matrix for 1971 Model Year Emissions	20
3	Stepwise Regression Results; Significant Independent Variables for 1971 Model Year Emissions	22
4	Significant Deterioration Factors by Model Years (All FY)	25
5	Significant Deterioration Differences: Model Year Effects (All FY)	26
6	Significant Deterioration Factors by EFP Fiscal Year (Model Year 1971)	27
7	Significant Deterioration Differences: Fiscal Year Effects (Model Year 1971)	29
8	Significant Deterioration Factors by EFP Fiscal Year (All Model Years)	30
9	Significant Deterioration Differences: Fiscal Year Effects (All Model Years)	31
10	Significant Deterioration Factors by Test Site (Model Year 1971)	32
11	Significant Deterioration Differences: Test Site Effects (Model Year 1971)	33
12	Significant Deterioration Factors by CID (Model Year 1971)	35
13	Significant Deterioration Differences: CID Effects (Model Year 1971)	36
14	Significant Deterioration Factors by Mileage Accumulation Rate (Model Year 1971)	37
15	Significant Deterioration Differences: Mileage Accumulation Rate Effects (Model Year 1971)	38
16	SDS Significant Deterioration Factors by Model Year for Low-Altitude, Non-California Cities	39
17	SDS Significant Deterioration Factors by Model Year for Denver	40

LIST OF TABLES (cont.)

Table	Title	Page
18	SDS Significant Deterioration Factors by Model Year for Los Angeles	41
19	SDS Significant Deterioration Differences: Model Year Effects (Low-Altitude, Non-California Cities)	43
20	SDS Significant Deterioration Differences: Test Site Effects (Model Years 1972-1974)	44
21	Significant Changes with Mileage in the Modal Emission Model Coefficient A1 (Constant Term)	46
22	Significant Changes with Mileage in the Modal Emission Model Coefficient A2 (Velocity Term)	47
23	Significant Changes with Mileage in the Modal Emission Model Coefficient A3 (Acceleration Term)	48
24	Comparison of SDS Bag Value Deterioration Factors with Modal Emissions Model Component Trends	49
25	SDS Significant Deterioration Factors in Idle Mode Bag Values by Test Site (Model Years 1972-1974)	51
26	SDS Idle Mode Significant Deterioration Differences: Test Site Effects (Model Years 1972-1974)	52
27	Prediction Error for 1971 Model Year Emissions	56
II-1	Regression Results by Model Year (EFP FY 71-73)	II-2
II-2	Regression Results by Fiscal Year (Model Year 1971)	II-4
II-3	Regression Results by Fiscal Year (All Model Years)	II-5
II-4	Regression Results for All Model Years (EFP FY 71-73)	II-6
II-5	Regression Results by Test Site	II-7
II-6	Regression Results by CID (EFP FY 71-73)	II-8
II-7	Regression Results by Accumulated Mileage Range (EFP FY 71-73)	II-9
II-8	Age and Mileage Regression Results (EFP FY 71-73)	II-10

LIST OF TABLES (cont.)

Table	Title	Page
III-1	Bag Value Regression Results for Pre-1968 Vehicles	III-2
III-2	Bag Value Regression Results for Model Years 1968 and 1969 . .	III-3
III-3	Bag Value Regression Results for Model Years 1970 and 1971 . .	III-4
III-4	Bag Value Regression Results for Model Years 1972-1974	III-5
III-5	Regression Results for Modal Emission Model Coefficient A1 .	III-6
III-6	Regression Results for Modal Emission Model Coefficient A2 .	III-7
III-7	Regression Results for Modal Emission Model Coefficient A3 .	III-8
III-8	Regression Results for Idle Mode Bag Values (EFP FY 73) . .	III-9

Section 1

INTRODUCTION

1.1 BACKGROUND

The Emission Factor Programs (EFP) are a series of projects, conducted annually, to measure exhaust emissions of in-use motor vehicles tested in as-received condition. The purpose of the EFP is to estimate total emissions of in-use vehicles by model year and region for the current calendar year. The results of the programs have also been used to develop linear emission deterioration curves based upon accumulated mileage. The deterioration curves have been used for projecting average emissions in order to assess how well vehicles are expected to perform in relation to the Federal standards over their useful life.

Accurate determination of emission deterioration levels has recently become more critical with the delay of new and more stringent standards. Moreover, the potential benefits of more stringent Federal standards, as opposed to implementation of Inspection/Maintenance, is in large part dependent on the performance of in-use vehicles. The important policy questions which depend on an assessment of the effectiveness of current Federal standards suggest that additional and thorough investigations of available EFP data be undertaken in order to develop the best emission deterioration estimates and to determine if these estimates can be expected to provide a meaningful and reliable assessment of vehicle emissions in successive years.

Modal emissions taken over the Surveillance Driving Sequence (SDS) are also collected in the EFPs in order to obtain a sound data base for the modal emission model. The model enables the user to predict emissions over any specified driving sequence. A precise estimate of modal emission deterioration is needed in order to accurately estimate the modal emissions of older vehicles which may not be tested over the SDS in successive EFPs.

1.2 PURPOSE AND GENERAL APPROACH

The purpose of this study is to estimate emission deterioration of in-use vehicles as they accumulate age and mileage. The analysis uses data collected on vehicles tested in the FY 71, FY 72, and FY 73 EFPs. Deterioration estimates are made using emissions measured from the Federal Test Procedure (FTP) and the SDS. The estimates are used to assess differences in deterioration rates between --

a. Model years, over the three fiscal years
b. Fiscal years, for a specific model year
c. Fiscal years, over all model years
d. Test sites (Denver, Los Angeles, and low altitude, non-California cities), for a specific model year.

This study includes investigations to determine the best form of the emission versus age and/or mileage curve (linear, quadratic, etc.) and to identify the most appropriate independent variable, age or mileage, or combinations, on which the emission deterioration rate estimates should be based. Also included is the identification of other vehicle and driving parameters (CID, mileage accumulation rate, etc.) which may affect deterioration.

This study also investigates changes in the SDS data modal emissions model coefficients as related to accumulated mileage. These coefficients have previously been determined for each vehicle in the FY 71 and FY 73 EFPs (see References 1 and 2).

In addition to the above specific investigations, this report also reviews the problems of predicting emissions with available EFP data, including the statistical and engineering significance of the predictions, the danger of extrapolating beyond age and/or mileage ranges, and the shortcomings of the EFP design and format for predictive purposes.

Section 2 of this report summarizes the study findings. Section 3 describes the emission deterioration analysis procedures and discusses the choice of deterioration curve form and independent variable. Deterioration analysis results are presented in Section 4. Problems and recommendations are included in Section 5.

Section 2

SUMMARY

The results of this study are summarized below in terms of general observations and conclusions regarding the appropriate form of the deterioration curves and the choice of an independent variable, specific deterioration analysis results, and a review of problems and recommendations. The results apply only to the test data generated in the FY 71-73 Emission Factor Programs. Extrapolation of data trends or application to newer vehicles may not be valid. In much of the analyses, Model Year 1971 vehicles have been used as a data base. These vehicles are the latest models to have been tested in all three EFPs.

2.1 CHOICE OF DETERIORATION CURVE FORM AND INDEPENDENT VARIABLE

1) For regressions of emission rate vs. Mileage, linear functions are adequate; higher order polynomials are not necessary to model any of the three pollutant trends. For regressions of emission rate vs. Age, quadratic functions appear to be necessary, but this may be a result of the inherent error in assigning a single Age value to all vehicles from one model year.

2) Accumulated Mileage appears to be preferable to Age as the independent variable in deterioration analysis. Either parameter can be used, but since emission deterioration as a function of Mileage can be assumed to be linear, the analysis is less complex than for Age deterioration analysis. Furthermore, the mileage parameter allows distinguishing between vehicles of the same model year. A combination of Mileage and Age does not appear to be warranted for the EFP data.

3) CO and NOX emissions for Model Year 1971 are more highly correlated with Weight and CID than with either deterioration parameter. This correlation is expected to decrease as newer vehicles meet more stringent standards.

2.2 DETERIORATION ANALYSIS RESULTS -- FTP DATA

1) Model Year Effects -- Model Years 1970-72 have essentially indistinguishable deterioration rates with mileage. The HC and CO deterioration

rates of this group are greater than those of earlier models (pre-1968). Apparent differences between this group and 1973 Models are due to the limited data and the short mileage span represented by the data for 1973 Models in the FY 73 EFP. NOX deterioration is essentially insignificant for all Model Years. [Analysis of Model Years 1968 and 1969 was not made.

2) Fiscal Year Effects -- Data for 1971 vehicles shows significant HC and CO deteriorations in FY 73 EFP only. This trend is more indicative of the increased range of mileage represented in the test data than of actual changes in deterioration rates. The data for all model years combined does exhibit significant differences from one EFP to the next; but these differences are probably as much due to the changing model year population as they are to a change in deterioration rates of specific vehicle groups.

3) Overall Emission Deterioration -- The overall emission rate regressions for all models (pre-68 to 1973) over the three EFPs are listed below. The second term in each equation is the estimated deterioration rate for each pollutant:

$$HC = 3.01 \text{ (gm/mi)} + 0.666 \text{ (gm/mi) per 10,000 miles}$$

$$CO = 43.41 \text{ (gm/mi)} + 7.170 \text{ (gm/mi) per 10,000 miles}$$

$$NOXC = 4.41 \text{ (gm/mi)} - 0.029 \text{ (gm/mi) per 10,000 miles}$$

4) Test Site Effects (Model Year 1971) -- The only significant difference in emission rate deterioration is exhibited in CO emissions. The deterioration rate is statistically greater for vehicles tested at low altitude, non-California sites than for vehicles tested at either Denver or Los Angeles.

5) CID Effects (Model Year 1971) -- Statistically significant differences are shown for the deterioration rates of vehicles grouped by engine size. CO deterioration rates are greater for medium engines (250-360 CID) than for smaller or larger engines. Both HC and NOX deterioration rates are higher for small engines than for large engines.

6) Mileage Accumulation Rate Effects (Model Year 1971) -- CO deterioration rates for low annual mileage accumulation vehicles are lower than for the higher accumulation rate vehicles (10,000 miles/year and up). HC and NOX differences are also exhibited.

2.3 DETERIORATION ANALYSIS RESULTS -- SDS DATA

1) Observed Bag Values, Model Year Effects -- The results are, in general, consistent with the FTP data results.

2) Observed Bag Values, Test Site Effects (Model Years 1972-74) -- The statistically significant results are not at all consistent with those for Model Year 1971 from the FTP Data: HC and CO deterioration rates were highest for vehicles tested in Denver. The low altitude, non-California site vehicles had the lowest CO deterioration rates.

3) Modal Emission Model Coefficients (Model Years 1972-74) -- Statistically significant changes in model parameters a_1 (constant term), a_2 (velocity term), and a_3 (acceleration term) with mileage accumulation are evident. Analysis of these trends shows several patterns:

a. Wherever significant trends exist in both the a_1 and a_2 components, these trends are of opposite sign.

b. Wherever significant factors exist for both the bag value and a_1 component, the a_1 trend has the same sign as the HC and CO deteriorations and the opposite sign from the NOX deterioration.

c. CO deterioration does not show up as an a_3 (acceleration trend; CO deterioration is always exhibited as an a_1 (constant) trend.

d. Both bag value emission deterioration and emission model coefficient changes are less evident in the Los Angeles tested vehicles than in vehicles from other test sites.

4) Observed Idle Mode Bag Values, Test Site Effects (Model Years 1972-74) -- Deterioration rate differences are quite consistent with those observed for the total bag values (item (2) above).

2.4 PROBLEMS AND RECOMMENDATIONS

This study shows that the EFP data can be used to characterize the tested vehicles in terms of historic emission trends with mileage and age. The applicability of such characterizations to the overall population depends highly on the number and stratification of vehicles used in the test samples. Serious problems exist, moreover, if these historical trends are used to predict emission levels or if emission deterioration of individual vehicles is required.

Recommendations to address the prediction problems include --

1) Consideration of historical, current, and postulated vehicle mixes in terms of model years, engine types, emission control equipment, and expected accumulated mileage levels per year.

2) Consideration of the number of vehicles from a specific group which are required to be tested over a period of time to yield emission predictions of specified precision.

3) Consideration of the form of the emission deterioration curves for newer vehicles.

4) Comparison of emission level predictions derived from the FY 71-73 EFP data as contained in this study with actual emission levels measured at this time (1976).

The analysis and comparison of emission deterioration for individual vehicles requires additional tests or a redesign of the EFP test format. Deterioration rates must be determined for individual vehicles over a period of time. Since deterioration is a consequence of vehicle maintenance, it is necessary to monitor the deterioration of vehicles maintained under various inspection and service schedules. It will be necessary to trade off the costs of such a test program with the expected gains in information.

An additional advantage in measuring emission characteristics for individual vehicles is that it becomes possible to analyze the correlation among the deteriorations of the three pollutants. Just as there exists a correlation between emission levels at any one time, it is suspected that the rates of change in these levels over a period of time are also correlated. It may be possible to develop a theory of composite deterioration by applying Principal Component Analysis and Factor Analysis procedures to emission deterioration data from individual vehicles.

Section 3

DETERIORATION ANALYSIS

This section describes the data used in this study and the data analysis and statistical techniques employed. Two primary investigations--the choice of the deterioration curve form, and the choice of the independent variable--are also presented as necessary first steps, prior to the actual data analysis given in Section 4.

3.1 DATA

The exhaust emissions data used in this study was collected on vehicles tested in the FY 71, FY 72, and FY 73 Emission Factor Programs (EFP). References 3, 4 and 5 summarize the analysis of the test data from these EFPs. The test data collected includes total HC, CO, and NOX (corrected) emissions for vehicles driven through the 1975 Federal Test Procedure, in units of grams/mile. Vehicle parameters include make, model year, CID, inertia weight, mileage, and emission control type.

Similar emission data exist for vehicles driven over the first 1024 seconds of the Surveillance Driving Sequence (SDS) for the FY 71 and FY 73 EFPs. In addition, emissions are available mode-by-mode (37 modes in all) for each vehicle in the SDS. These modal emissions form the basis for a modal emission model (References 1 and 2). The coefficients used in the model for each vehicle--weighting values for each of the velocity/acceleration terms--are analyzed in this study to estimate changes in the coefficients as a function of time and/or mileage. Idle mode bag value deteriorations are also analyzed.

3.2 DATA ANALYSIS AND STATISTICAL TESTS

The regression analyses conducted in this study--linear, polynomial, stepwise linear--have employed the BMD Biomedical Computer Programs (Health Sciences Computing Facility, UCLA). These or similar statistical analysis routines are widely available and are not documented here.

Testing for the significance of a term in an emission deterioration equation involves an F-test (sometimes called a partial F-test) as described in Reference 6. This standard procedure is used, for example, to test for the significance of the quadratic term when investigating the form of the deterioration equation. The procedure is also used to test the hypothesis that a deterioration factor is not statistically different from zero. The $1 - \alpha$ level of significance for which the hypothesis can be rejected is a measure of how much of an age or mileage trend is exhibited in the emission data. The engineering significance of this trend is a matter of judgment.

Testing for statistically significant differences between two deterioration slopes (in the linear case), due to model year or test site for example, involves a special T-test as described in Reference 7. Again, the $1 - \alpha$ level of significance is a measure of the confidence with which one can reject the hypothesis that two deterioration rates are not statistically different. The direction of any difference can be ascertained from the two slope estimates. The magnitude of the difference (and its engineering significance), however, cannot be implied from the test. Appendix I briefly describes the test computations.

The method of least squares for obtaining estimates of slopes and intercepts does not depend upon any assumption about the distribution of emissions. Statements about the confidence in these estimates, however, require knowledge of (or assumptions about) the distribution of these estimates. This requirement in turn implies the need for specifying the distribution form of the emission data. Throughout this study the emission data has been assumed to have a normal distribution. Consequently, the intercept and slope estimates are also assumed to be normally distributed. Inasmuch as the emissions distributions often appear to have an asymmetric, e.g., log-normal, form, the assumption of normality can lead to errors in some of the results. A more exact method would have been to normalize the data by taking the log of the emissions data, for example. These considerations indicate, however, that the approximate method (assuming normality) yields sufficient accuracy for this study:

1) In all cases involving tests on the significance of deterioration factors, the data set has "outliers" removed. These outliers appear in the high emissions "tail" of the distribution. Their removal reduces the degree of asymmetry in the emissions distribution.

2) A relatively large number of sample points are used in any one regression analysis; so the resulting confidence intervals for intercept and slope estimates are small and concentrated near the mean where any asymmetry in the actual distribution has only a minor effect. As the confidence interval gets wider as a result of decreased sample sizes, the effect of the asymmetric tails becomes more important.

3) When establishing confidence limits on the sample mean emission level for a specified mileage or age (see Section 5), there is relief from the distribution form assumption in that the distribution of the sample mean can be considered to be normal. By the Central Limit Theorem the distribution of means of samples taken from <u>any</u> distribution approaches a normal distribution as the number of samples approaches infinity. The sample number which can be considered "sufficiently close to infinity" depends on the shape of the distribution of the original variables. Even when the underlying distribution is rectangular or triangular, the distribution of mean values from samples as small as <u>four</u> is approximately normal.

3.3 FORM OF THE DETERIORATION CURVES

Deterioration curves have commonly been assumed to be linear with age or mileage, in some cases due to lack of sufficient data to justify more complex functions, and in other cases merely in order to avoid the complexities of determining and analyzing a nonlinear function. The choice of the form of the curves is, of course, closely tied to the choice of the independent variable. In this section, it is assumed that either <u>age</u> or <u>mileage</u> alone will be the choice for the independent variable. Further refinement of this assumption is presented in Section 3.4.

The procedure for investigating the form of the deterioration curves involves a combination of statistical and practical considerations. First, the FTP emissions data from all three EFPs for 1971 model year vehicles in low altitude, non-California was chosen as the data base for

the analysis. This data set is, therefore, based on the latest vehicle model year which has been tested in all three EFPs. The three year span and associated mileage range for the vehicles is necessary to point out possible nonlinear deterioration effects. The data set did not include data collected at the Denver or Los Angeles test sites in order to avoid any confounding with altitude effects or California emission control equipment differences.

The data was first analyzed assuming linear models:

$$\text{HC, CO, or NOXC} = a_i + b_i \times \text{MILEAGE(miles)}.$$

and

$$\text{HC, CO, or NOXC} = a_j + b_j \times \text{AGE(months)}.$$

Figures 1 through 6 indicate the data points and the linear regressions (solid lines) for the six cases. The age values (in months) for vehicles in Figures 4 through 6 are approximated by considering the time over which the EFP was conducted and the average production date for vehicles of the specified model year (1971). True age--in terms of day of production or sale to day of testing for each vehicle--was not available.

Also shown on Figures 1 through 6 are the regression results (dashed curves) when quadratic models are assumed:

$$\text{HC, CO, or NOXC} = a_i + b_i \times \text{MILEAGE} + c_i \times (\text{MILEAGE})^2$$

and

$$\text{HC, CO, or NOXC} = a_j + b_j \times \text{AGE} + c_j \times (\text{AGE})^2$$

From these figures it is apparent that in most cases the quadratic model is unwarranted. Table 1 gives the results of F-tests which indicate the relative significance of the terms in the two models. Consistent with the plots, the statistical tests show that none of the pollutant deteriorations as functions of mileage warrant a quadratic term. As functions of age, all of the pollutant deteriorations exhibit some nonlinear effects--NOXC quite significantly (even to the exclusion of the linear term); HC and CO moderately (the linear term is more significant, in a statistical sense, than the quadratic term). It must be emphasized, however, that these nonlinear effects are more likely to be a result of the error in assigning a specific age to all vehicles of one model year than they are a result of actual deterioration rate patterns.

Figure 1 LINEAR AND QUADRATIC HC VS. MILEAGE REGRESSIONS FOR 1971 MODEL YEAR DATA

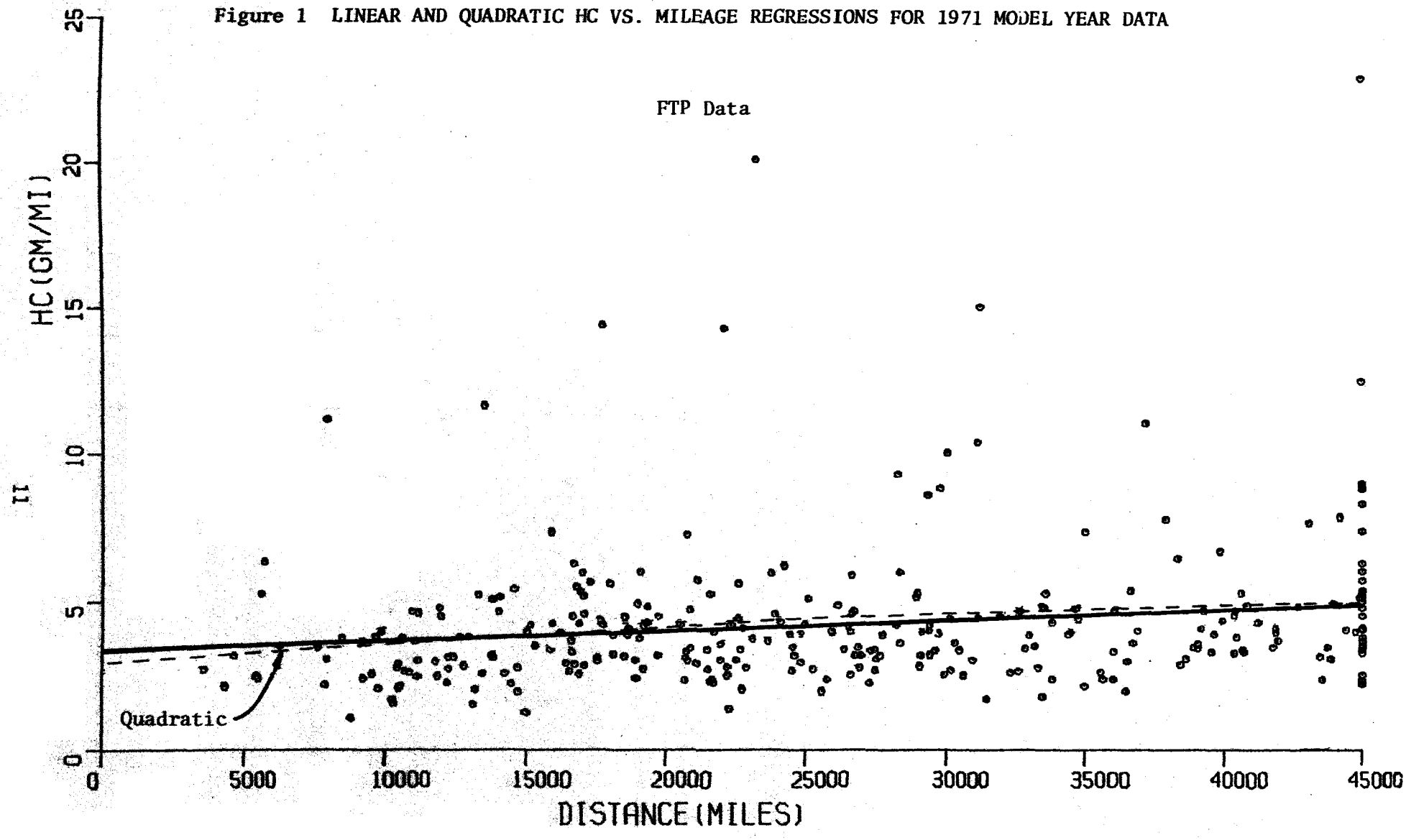


Figure 2 LINEAR AND QUADRATIC CO VS. MILEAGE REGRESSIONS FOR 1971 MODEL YEAR DATA

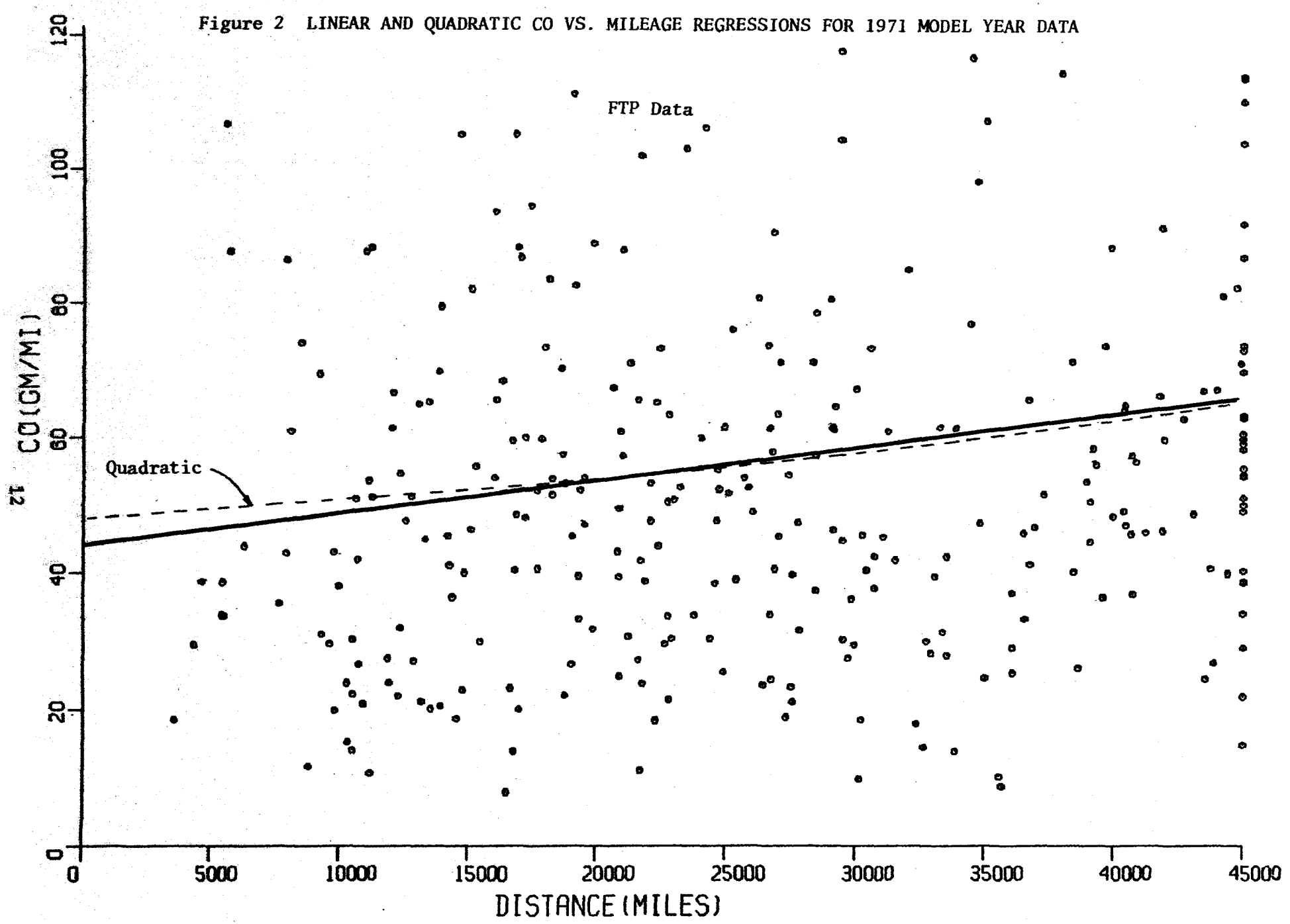
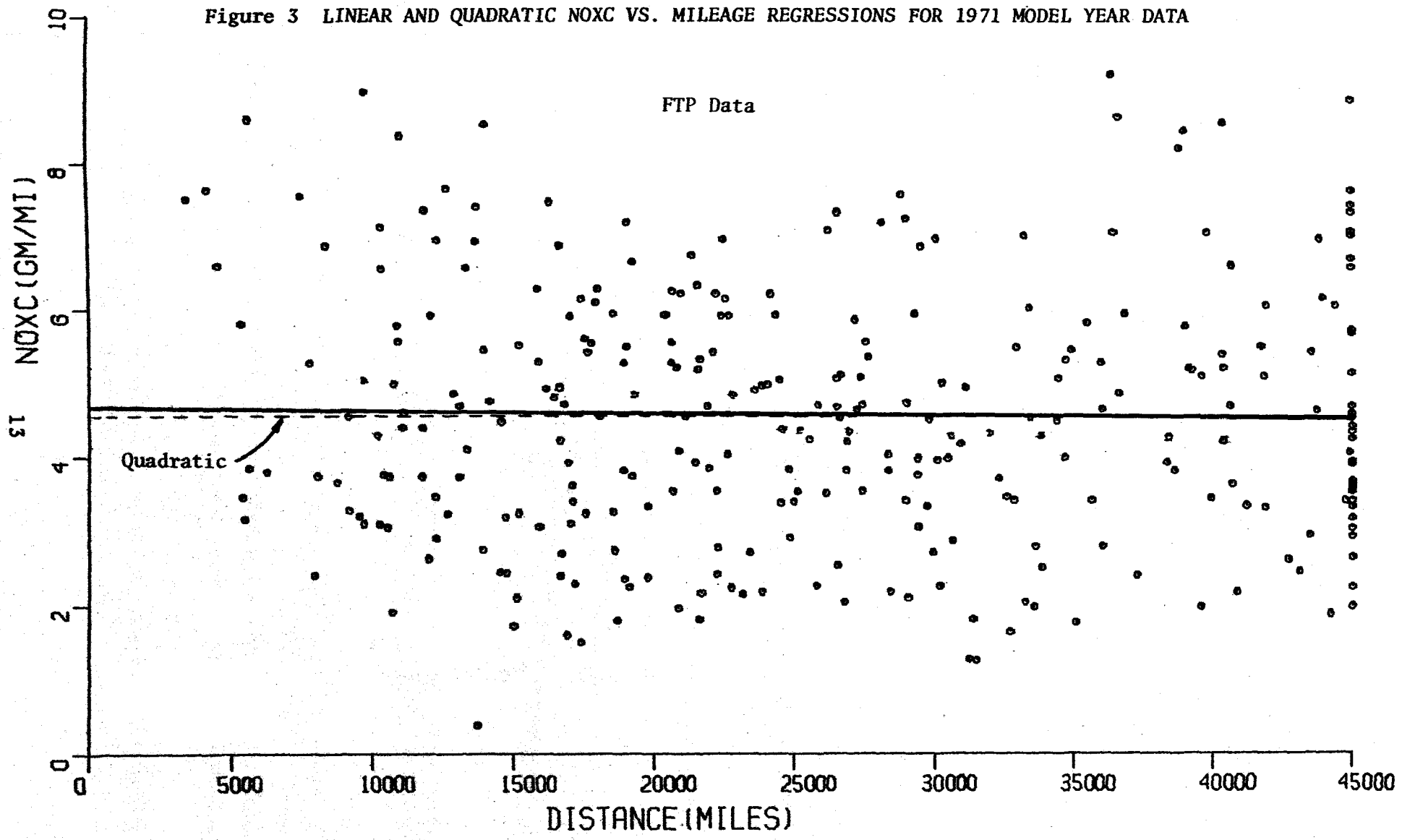


Figure 3 LINEAR AND QUADRATIC NOXC VS. MILEAGE REGRESSIONS FOR 1971 MODEL YEAR DATA



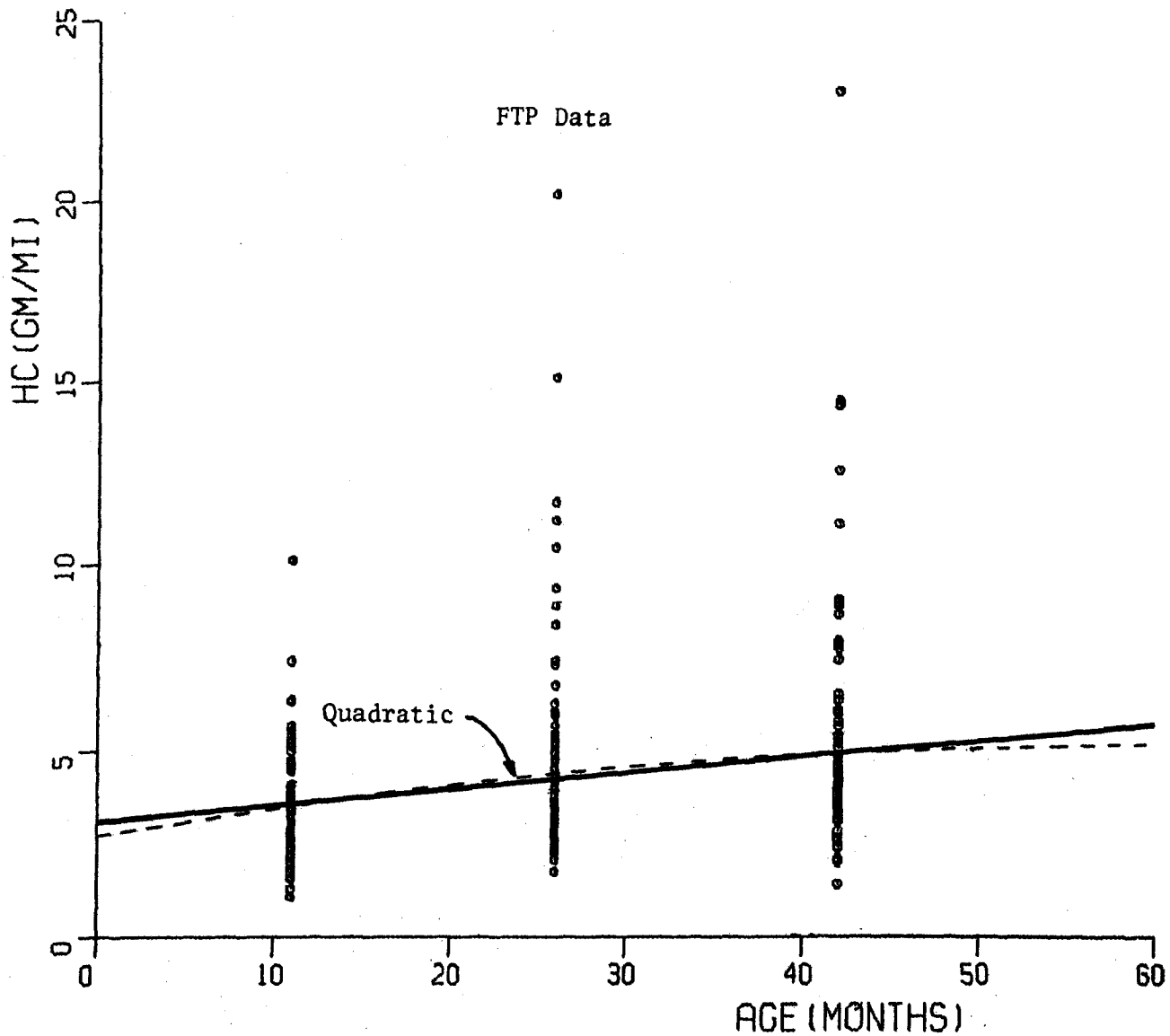


Figure 4 LINEAR AND QUADRATIC HC VS. AGE REGRESSIONS FOR 1971 MODEL YEAR DATA

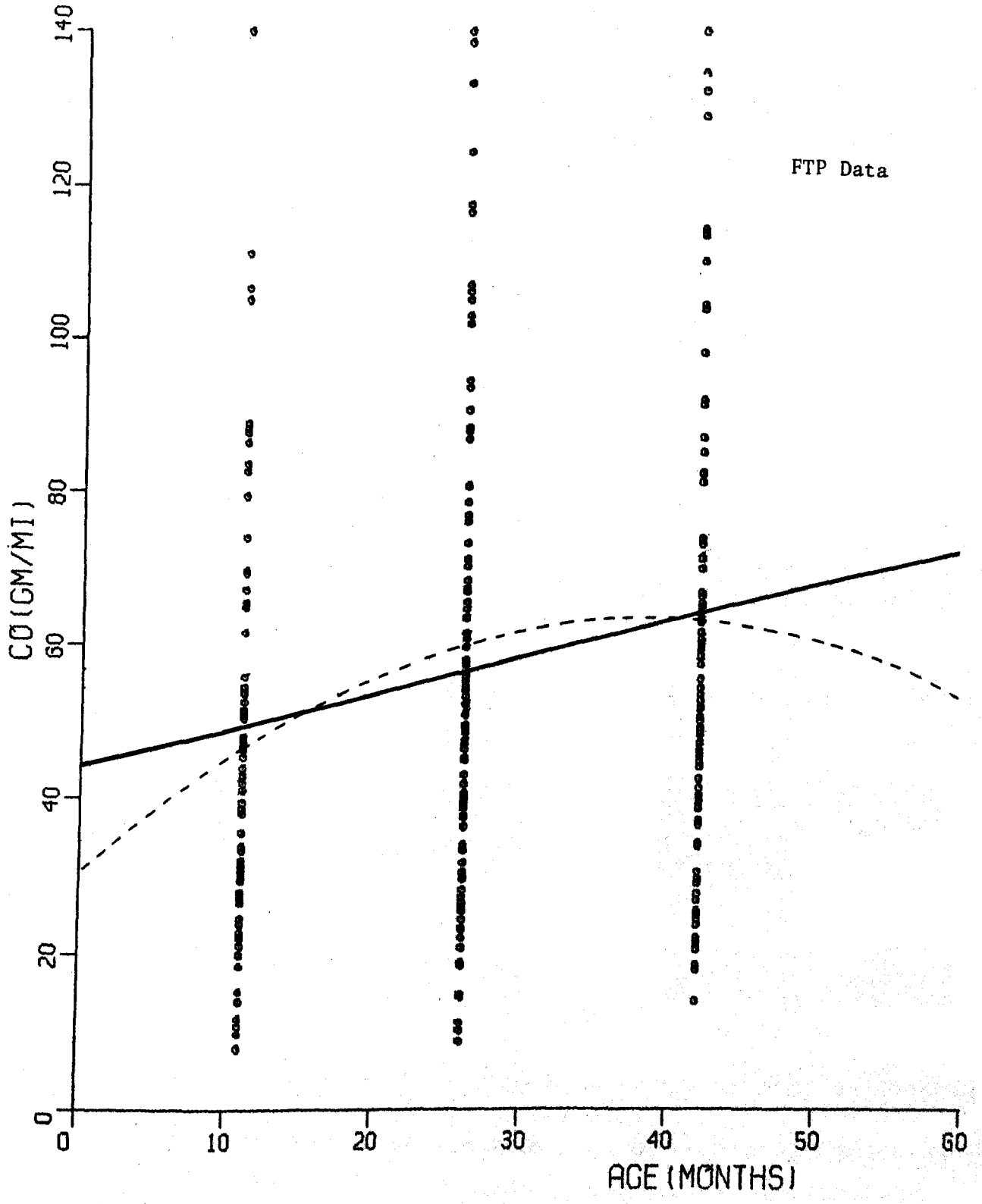


Figure 5 LINEAR AND QUADRATIC CO VS. AGE REGRESSIONS FOR 1971 MODEL YEAR DATA

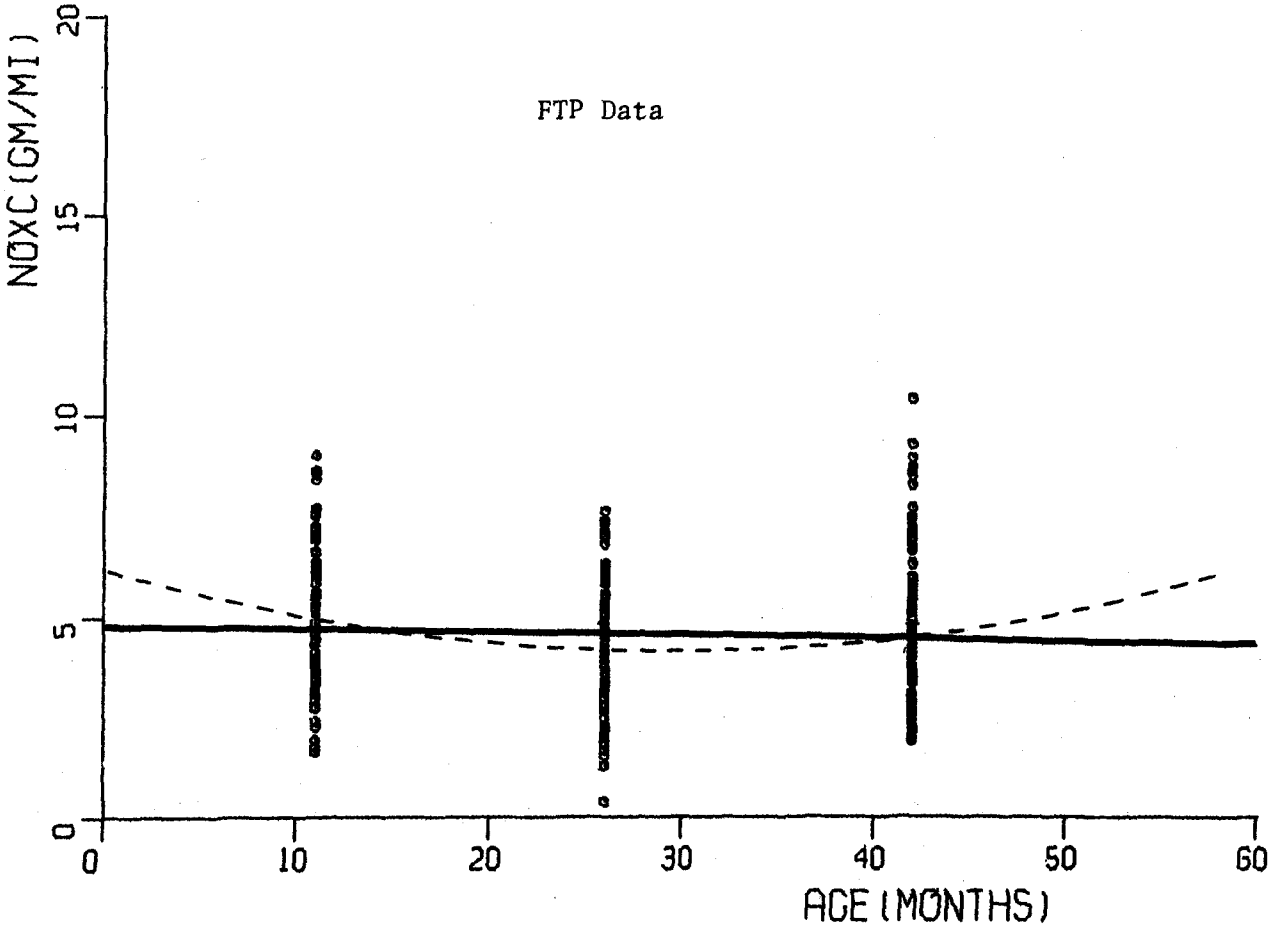


Figure 6 LINEAR AND QUADRATIC NOXC VS. AGE REGRESSIONS FOR 1971 MODEL YEAR DATA

Table 1

COMPARISON OF LINEAR AND QUADRATIC MODELS
OF EMISSION DETERIORATION

FTP Data Model Year 1971, EFP FY 71-73
Low Altitude Non-California Cities

	MODEL	TERM	POLLUTANT	F-TEST RATIO	F-TEST SIGNIFICANCE LEVEL
MILEAGE FUNCTIONS	LINEAR	LINEAR	HC	12.89	.999
			CO	11.63	.999
			NOXC	0.20	---
	QUADRATIC	LINEAR	HC	12.89	.999
			CO	11.61	.999
			NOXC	.20	---
		QUAD	HC	1.12	---
			CO	0.49	---
			NOXC	0.22	---
AGE FUNCTIONS	LINEAR	LINEAR	HC	15.12	.999
			CO	7.88	.99
			NOXC	.80	---
	QUADRATIC	LINEAR	HC	15.12	.999
			CO	7.92	.99
			NOXC	.82	---
		QUAD	HC	1.05	.75
			CO	2.47	.75
			NOXC	8.55	.995

An extension of this investigation to higher order polynomials or more complex functions is unnecessary--it is not warranted with the mileage functions, and it is impossible with the age functions since a maximum of only three values of the independent variable for any one model year are available in this study.

3.4 INDEPENDENT VARIABLE INVESTIGATION

Emission deterioration can be a function of vehicle mileage, as a result of engine wear, for example; and it can be a function of vehicle age, as a result of corrosive surface aging, for example. One of the investigations conducted in this study considered which variable, mileage or age (or a combination of the two), is the more appropriate independent variable (explains more of the variability about the deterioration curve) within the EFP emission data.

A stepwise linear regression program was used as a "shotgun" approach for comparing the contributions of the age and mileage parameters, along with several others, toward accounting for the overall variability exhibited in the emissions data. Once again the 1971 Model Year data from the three EFPs was used. The variables allowed to "enter" the stepwise regression process for each pollutant were:

- Age
- Mileage
- CID
- Weight
- Make Code
- Mileage/CID
- Mileage/Age
- Mileage/Weight

Some of these variables were not expected to emerge as important. The "Make Code," for example, is a rather arbitrary integer coding used for data identification purposes only. The "Mileage over CID" and "Mileage over Weight" ratios are also somewhat arbitrary, unless one suspected a strong correlation between miles driven and the size or type of vehicle. The

"Mileage over Age" ratio, on the other hand, carries the implication of "mileage accumulation rate" which, indeed, might be a secondary factor in emission deterioration.

The first output of the stepwise regression program is the correlation matrix which identifies the correlation between each of the eleven variables (the eight listed above, plus the three pollutants). The important patterns which emerged are listed below; the complete correlation matrix is given in Table 2.

1. HC emissions are most highly correlated with CO emissions (0.496); then with AGE (0.217) and MILEAGE (0.201).
2. CO emissions are most highly correlated with HC emissions (0.496); then with CID (0.283) and WEIGHT (0.228); then with MILEAGE (0.191) and AGE (0.158).
3. NOXC emissions are most highly correlated with WEIGHT (0.367) and CID (0.319); NOXC emissions are negatively correlated with CO emissions (-.224) and HC emissions (-.131); they are essentially uncorrelated with AGE (-.051) and MILEAGE (-.026).

Without even going to the stepwise regression results, several important, albeit not new, observations can be made:

A. The emission levels of the pollutants are closely correlated to each other. This implies that a mechanism to analyze the deterioration rates of all three pollutants simultaneously, e.g. Factor Analysis, could yield benefits beyond those attained by analyzing each pollutant deterioration separately.

B. AGE and MILEAGE (closely correlated to each other) are approximately equal in importance as independent variables.

Table 2

CORRELATION MATRIX FOR 1971 MODEL YEAR EMISSIONS

FTP DATA, EFP FY 71-73 Low Altitude, Non-California Cities

Variable Name	Variable Name
1. MAKE	7. CO
2. CID	8. NOXC
3. AGE	9. MILEAGE/CID
4. WEIGHT	10. MILEAGE/AGE
5. MILEAGE	11. MILEAGE/WEIGHT
6. HC	

Variable	1	2	3	4	5	6	7	8	9	10	11
1	1.000	-.209	0.050	-.217	0.026	-.021	-.055	-.024	0.230	-.013	0.169
2		1.000	0.014	0.939	0.115	0.044	0.283	0.319	-.640	0.079	-.389
3			1.000	-.030	0.594	0.217	0.158	-.051	0.317	-.369	0.527
4				1.000	0.096	0.008	0.228	0.367	-.583	0.107	-.412
5					1.000	0.201	0.191	-.026	0.481	0.413	0.825
6						1.000	0.496	-.131	0.120	-.014	0.200
7							1.000	-.224	-.103	-.008	0.030
8								1.000	-.281	0.044	-.227
9									1.000	0.179	0.847
10										1.000	0.316
11											1.000

C. CID and WEIGHT (closely correlated to each other) account for more of the CO and NOXC emissions variability than do either AGE or MILEAGE. This implies that any changes in the sampling distribution of vehicles by weight or CID from one test year to the next could easily confound the analysis of emission changes with age or mileage. (This correlation should diminish for newer vehicles subject to more stringent controls.)

The stepwise regression output lists the order in which the variables "enter" the linear model, most important first. Table 3 shows those variables which were "significant" in explaining emission variability for each pollutant. In the stepwise regression process, the pollutants themselves were restricted to be only dependent variables; i.e., they could not enter the regression of another pollutant. The resulting lists in Table 3 indicate that MILEAGE is the overall better choice for an independent variable to describe emission deterioration. A combined function of AGE and MILEAGE does not seem warranted. Even in the HC emissions where both variables are significant, the first variable to enter the regression is more important by a factor of about 5 to 1. MILEAGE did not enter first in this case since it was slightly less correlated than AGE with HC emissions.

Other considerations point to MILEAGE as the better choice. For those late Model Years where testing in only one EFP has occurred, AGE would be useless as a variable since all vehicles of one Model Year have nominally the same age. On the other hand, AGE is a much easier parameter to use when characterizing a vehicle population. The cost of this simplicity is apparently, however, a loss of some precision in estimating emission deterioration. Furthermore, from Section 3.3 it is evident that nonlinear AGE deterioration curves may be necessary (or the associated error in the AGE value accounted for) when estimating emission deteriorations; whereas, linear MILEAGE deterioration curves are adequate for all three pollutants.

Table 3

STEPWISE REGRESSION RESULTS; SIGNIFICANT INDEPENDENT VARIABLES FOR 1971 MODEL YEAR EMISSIONS

FTP DATA EFP FY 71-73 Low Altitude, Non-California Cities

HC Emissions	CO Emissions	NOXC Emissions
1. AGE	1. CID	1. WEIGHT
2. MILEAGE	2. MILEAGE	2. MILEAGE
3. MILEAGE/WEIGHT	3. MILEAGE/AGE	3. MILEAGE/WEIGHT
4. CID	4. WEIGHT	4. CID

Section 4

DETERIORATION ANALYSIS RESULTS

This section discusses the results of the deterioration analyses conducted on the FTP and SDS data from the FY 71 - 73 Emission Factor Programs (EFP). The analyses are based on linear regressions of emissions vs. mileage, in keeping with the results of Section 3. Thus the term deterioration factor or slope is appropriate.

The regressions were conducted in two ways. First the regressions were performed using the complete basic set of vehicles as listed on the EFP data tapes. Second, the regressions were performed on a "cleaned up" data set which remained after "outlier" vehicles had been eliminated. "Outliers" in this case are any vehicles whose vector of HC, CO and NOXC emission levels falls outside of the 1.65σ (90%) ellipsoid. Thus, any vehicle whose HC emissions are greater than $1.65\sigma_{HC}$ is eliminated. Moreover, a vehicle which has two or three high levels, although all within the individual 1.65σ limits, might also be eliminated if its 3-dimensional emission vector falls outside the 1.65σ joint distribution emission ellipsoid. Outliers were found to occur throughout the mileage range -- not just at the high mileages.

The data analysis is presented in this section for the corrected, or cleaned up, data since it better represents the expected deterioration effects in the whole population. The regression results for both the basic and corrected data are presented in Appendices II and III.

Tests for significance of deterioration factors and for differences between two factors use level of significance cutoffs of .75 and .80, respectively. These cutoffs have been chosen somewhat low to assure that if there is indeed a finite difference, it will be detected. Had the cutoff level been chosen higher, say at .99, fewer significant differences would have been noted. The cutoff values are slightly different (.75 and .80) as a result of the available F- and t-tables.

4.1 FTP DATA ANALYSIS

The chart below indicates the EFP Fiscal Year, Model Year, and Test Site data groups which were used in the regression analysis. X's indicate that the particular model year did not exist for that EFP Fiscal Year.

Additional regressions were also conducted to evaluate secondary effects due to CID and mileage accumulation rate.

MODEL YEAR

EFP FY	Pre '68	68, 69	70	71	72	73	All Available
71	---	---	---	Low			Low
72	---	---	---	Low	(Low)		Low
73	---	---	----	Low		(Low)	Low
All Available	Low	---	Low	Low, LA, Den	Low	Low	Low

Low = Low Altitude, Non-California LA = Los Angeles Den = Denver

Deterioration factor data in tables which follow is presented in scientific notation with the power of ten listed after the four-place characteristic.

4.1.1 Model Year Effects

Table 4 shows the deterioration factors which are significantly different from zero for the Model Year groups listed. Data from every applicable EFP Fiscal Year was used. The fact that Model Year 1973 has no significant deterioration factors is more an indication of the short mileage span represented by the data (EFP FY 73 only) than of the absence of emission deterioration.

Table 5 indicates the significant differences in deterioration factors between model years. Model Years 1970-72 have essentially indistinguishable factors. Each exhibits greater HC and CO deterioration than Model Year 1973, again, due to the lack of significance in the 1973 data.

4.1.2 EFP Fiscal Year Effects for a Given Model Year

In the section above it was shown that over a three year-period Model Year 1971 vehicles exhibited (statistically) significant deterioration rates in HC and CO (see Table 4). Table 6 shows the results when

Table 4

SIGNIFICANT DETERIORATION FACTORS
BY MODEL YEARS (ALL FY)

FTP Data
Low-Altitude, Non-California Cities

MODEL YEAR	Deterioration Factors Significantly Different from Zero		
	Pollutant	Deterioration Factor (Gm/Mi per Mile)	Significance Level
Pre-1968	HC	0.1486 -04	.90
	CO	0.2754 -03	.975
1970	HC	0.2267 -04	.999
	CO	0.4332 -03	.999
	NOXC	-.8343 -05	.75
1971	HC	0.3018 -04	.999
	CO	0.5137 -03	.999
1972 (FY 72 & 73)	HC	0.2536 -04	.999
	CO	0.4710 -03	.999
1973 (FY 73)	None	--	--

Table 5

SIGNIFICANT DETERIORATION DIFFERENCES:
MODEL YEAR EFFECTS (ALL FY)

FTP Data
Low-Altitude, Non-California Cities

Model Year Test	Significant Differences Between Deterioration Factors: Pollutant Deterioration Difference	Significance Level
1971 vs. Pre-1968	$HC_{71} > HC_{Pre-68}$ $CO_{71} > CO_{Pre-68}$	.80 .80
1971 vs. 1970	No significant differences	--
1971 vs. 1972	No significant differences	--
1971 vs. 1973	$HC_{71} > HC_{73}$ $CO_{71} > CO_{73}$	.98 .98
1972 vs. 1973	$HC_{72} > HC_{73}$ $CO_{72} > CO_{73}$	.90 .95
1970 vs. 1973	$HC_{70} > HC_{73}$ $CO_{70} > CO_{73}$	.80 .95

Table 6

SIGNIFICANT DETERIORATION FACTORS
BY EFP FISCAL YEAR (MODEL YEAR 1971)

FTP Data

Low-Altitude, Non-California Cities

EFP FISCAL YEAR	Deterioration Factors Significantly Different from Zero		
	Pollutant	Deterioration Factor (Gm/Mi per Mile)	Significance Level
1971	None	--	--
1972	None	--	--
1973	HC	0.2280 -04	.975
	CO	0.8169 -03	.999

the data for only one EFP Fiscal Year is used at a time. For FY 71 and 72 no significant factors appear. This result does not necessarily imply that the deterioration rate for Model Year 1971 vehicles suddenly increased in 1973 after two initial years with no significant deterioration. In fact, such a nonlinear effect has already been discounted in Section 3.3 for this same three year data set.

Table 7 indicates the significant differences in deterioration between the EFP Fiscal Years for Model Year 1971. Only CO deterioration shows differences at a level greater than 0.75.

4.1.3 EFP Fiscal Year Effects for All Model Years

Table 8 lists the significant deterioration rates for all vehicles by EFP Fiscal Year. Also shown are the results for the total data set -- all Fiscal Years and all Model Years. The large data sets provide significant factors in every case. It is not clear why the 73 EFP shows a positive NOXC deterioration factor.

Table 9 indicates that NOX deterioration rates are increasing from year to year (from negative to positive, in fact) over all Model Years. HC and CO deterioration rates show the opposite trend. This pattern is probably due as much to the changing Model Year population from one fiscal year to the next as it is to a change in the deterioration rates of the constant elements of the population. The fact that emission standards also changed in some years implies that these "overall" deterioration rates are somewhat meaningless.

4.1.4 Test Site Effects

Table 10 lists the significant deterioration factors for Denver, Los Angeles, and the other low altitude, non-California city test sites.

Table 11 shows that the only significant difference occurs with CO deterioration. It is statistically greater in the low, non-California cities than in either Denver or Los Angeles.

4.1.5 CID Effects

The fact that CO and NOX emissions were highly correlated to CID in the stepwise regression analysis (see Section 3.4) prompted the

Table 7

SIGNIFICANT DETERIORATION DIFFERENCES:
FISCAL YEAR EFFECTS (MODEL YEAR 1971)

FTP Data
Low-Altitude, Non-California Cities

EFP Fiscal Year Test	Significant Differences Between Deterioration Factors	
	Pollutant Deterioration Difference	Significance Level
FY 71 vs. FY 72	No significant differences	--
FY 71 vs. FY 73	$CO_{71} < CO_{73}$	.90
FY 72 vs. FY 73	$CO_{72} < CO_{73}$	.98

Table 8

SIGNIFICANT DETERIORATION FACTORS
BY EFP FISCAL YEAR (ALL MODEL YEARS)

FTP Data
Low-Altitude, Non-California Cities

EFP FISCAL YEAR	Deterioration Factors Significantly Different from Zero		
	Pollutant	Deterioration Factor (Gm/Mi per Mile)	Significance Level
1971	HC	0.6542 -04	.999
	CO	0.7634 -03	.999
	NOXC	-.2066 -04	.999
1972	HC	0.5899 -04	.999
	CO	0.6868 -03	.999
	NOXC	-.1183 -04	.999
1973	HC	0.3585 -04	.999
	CO	0.5909 -03	.999
	NOXC	0.1542 -04	.999
1971 - 1973	HC	0.5243 -04	.999
	CO	0.6678 -03	.999
	NOXC	-.2870 -05	.90

Table 9

SIGNIFICANT DETERIORATION DIFFERENCES:
FISCAL YEAR EFFECTS (ALL MODEL YEARS)

FTP Data
Low-Altitude, Non-California Cities

EFP Fiscal Year Test	Significant Differences Between Deterioration Factors	
	Pollutant Deterioration Difference	Significance Level
FY 71 vs. FY 72	$NOXC_{71} < NOXC_{72}$	.90
FY 71 vs. FY 73	$HC_{71} > HC_{73}$	.999
	$CO_{71} > CO_{73}$	.90
	$NOXC_{71} < NOXC_{73}$	.999
FY 72 vs. FY 73	$HC_{72} > HC_{73}$	.999
	$CO_{72} > CO_{73}$	.80
	$NOXC_{72} < NOXC_{73}$	.999

Table 10

SIGNIFICANT DETERIORATION FACTORS
BY TEST SITE (MODEL YEAR 1971)

FTP Data
EFP FY 71-73

TEST SITE	Deterioration Factors Significantly Different from Zero		
	Pollutant	Deterioration Factor (Gm/Mi per Mile)	Significance Level
Low Altitude, Non-Calif. Cities	HC	0.3018 -04	.999
	CO	0.5137 -03	.999
Denver	None	--	--
Los Angeles	HC	0.1958 -04	.975
	CO	0.2180 -03	.75
	NOXC	-.1411 -04	.95

Table 11

SIGNIFICANT DETERIORATION DIFFERENCES:
TEST SITE EFFECTS (MODEL YEAR 1971)

FTP Data
EFP FY 71-73
Low-Altitude, Non-California Cities

Test Site Test	Significant Differences Between Deterioration Factors	
	Pollutant Deterioration Difference	Significance Level
Low Altitude, Non-Calif. vs. Denver	$CO_{Low} > CO_{Denver}$	.95
Low Altitude, Non-Calif. vs. Los Angeles	$CO_{Low} > CO_{LA}$	.80
Denver vs. Los Angeles	No significant differences	--

suspicion that deterioration rates might also be affected by CID. Table 12 indicates that 1971 Model vehicles tested over all three EFPs, indeed, have significant deterioration rates when broken down by CID.

Table 13 shows that the CO deterioration rate is (statistically) significantly greater for the medium engines (250-360 CID) than for the small or large engines. Both HC and NOXC deterioration rates are higher for the small engines than for the large engines.

4.1.6 Mileage Accumulation Rate Effects

The rate at which mileage is accumulated is likely to be a factor in emission deterioration. Many short trips for example can result in greater engine wear than several long trips. Table 14 shows the significant deterioration factors for the 1971 Model vehicles (all three EFPs) broken down by an overall miles per year range. The three ranges were chosen to provide approximately equal sample sizes. The statistical significance between these factors is indicated in Table 15. CO deterioration rates for the low accumulation vehicles are lower than for the medium (up to 14,000 miles/year) or the high accumulation (14,000 to 100,000 miles/year) vehicles. Deterioration factors for the medium rate group are significantly higher than the other two groups for both HC and NOX emissions. The reasons behind these patterns are not evident.

4.2 SDS DATA ANALYSIS

The Surveillance Driving Sequence (SDS) data from EFP FY 71 and FY 73 includes observed bag values of emissions after the first 1024 seconds of the SDS and mode-by-mode emission values. These modal data form the basis for a modal emissions model (References 1 and 2). The coefficients used in the model for each vehicle -- weighting values for each of the velocity/acceleration terms -- are analyzed in this section to estimate changes in the model coefficients as a function of mileage. Deterioration factors are presented in scientific notation with the power of ten listed after the four-place characteristic.

4.2.1 Observed Bag Values -- Model Year and Test Site Effects

Tables 16-18 list the statistically significant deterioration factors by Model Year groups for data from the three test site categories

Table 12

SIGNIFICANT DETERIORATION FACTORS
BY CID (MODEL YEAR 1971)

FTP Data
EFP FY 71-73
Low-Altitude, Non-California Cities

CID RANGE	Deterioration Factors Significantly Different from Zero		
	Pollutant	Deterioration Factor (Gm/Mi per Mile)	Significance Level
0 - 249	HC	0.4312 -04	.995
	NOXC	-.4295 -04	.99
250 - 360	HC	0.2533 -04	.999
	CO	0.6518 -03	.999
361 - 500	HC	0.1510 -04	.75

Table 13

SIGNIFICANT DETERIORATION DIFFERENCES

CID EFFECTS (MODEL YEAR 1971)

FTP Data

EFP FY 71-73

Low-Altitude, Non-California Cities

CID Range Test	Significant Differences Between Deterioration Factors	
	Pollutant Deterioration Difference	Significance Level
Small vs. Medium	$CO_{Small} < CO_{Med}$	.95
Small vs. Large	$HC_{Small} > HC_{Large}$	.80
	$NOXC_{Small} > NOXC_{Large}$	.95
Medium vs. Large	$CO_{Med} > CO_{Large}$	.90

Table 14

SIGNIFICANT DETERIORATION FACTORS
BY MILEAGE ACCUMULATION RATE (MODEL YEAR 1971)

FTP Data
EFP FY 71-73
Low-Altitude, Non-California Cities

MILEAGE ACCUMULATION RATE (Miles per Year)	Deterioration Factors Significantly Different from Zero		
	Pollutant	Deterioration Factor (Gm/Mi per Mile)	Significance Level
0 - 9,999	HC	0.3768 -04	.975
	NOXC	-.2751 -04	.75
10,000 - 13,999	HC	0.5100 -04	.999
	CO	0.7030 -03	.999
	NOXC	0.1754 -04	.75
14,000 - 100,000	HC	0.2051 -04	.95
	CO	0.5820 -03	.995
	NOXC	-.1340 -04	.75

Table 15

SIGNIFICANT DETERIORATION DIFFERENCES:
MILEAGE ACCUMULATION RATE EFFECTS (MODEL YEAR 1971)

FTP Data
EFP FY 71-73
Low-Altitude, Non-California Cities

Mileage Accumulation Rate Test	Significant Differences Between Deterioration Factors	
	Pollutant Deterioration Difference	Significance Level
Low vs. Medium	$CO_{Low} < CO_{Med}$	.80
	$NOXC_{Low} < NOXC_{Med}$	.90
Low vs. High	$CO_{Low} < CO_{High}$	.80
Medium vs. High	$HC_{Med} > HC_{High}$	.95
	$NOXC_{Med} > NOXC_{High}$	.90

Table 16

SDS SIGNIFICANT DETERIORATION FACTORS
BY MODEL YEAR FOR LOW-ALTITUDE, NON-CALIFORNIA CITIES

MODEL YEARS	Deterioration Factors Significantly Different from Zero		
	Pollutant	Deterioration Factor (Gm/Mi per Mile)	Significance Level
Pre-1968	NOXC	-.7901 -05	.90
1968, 1969	HC	0.1621 -04	.95
	CO	0.2632 -03	.95
1970, 1971	HC	0.1974 -04	.995
	CO	0.1679 -03	.75
1972 - 1974	CO	-.2411 -03	.95
	NOXC	0.4119 -04	.999

Table 17

SDS SIGNIFICANT DETERIORATION FACTORS
BY MODEL YEAR FOR DENVER

MODEL YEARS	Deterioration Factors Significantly Different from Zero		
	Pollutant	Deterioration Factor (Gm/Mi per Mile)	Significance Level
Pre-1968	None	--	--
1968, 1969	NOXC	0.1888 -04	.75
1970, 1971	CO	-.8214 -03	.90
1972 - 1974	HC	0.4890 -04	.99
	CO	0.1020 -02	.75
	NOXC	0.3875 -04	.995

Table 18

SDS SIGNIFICANT DETERIORATION FACTORS
BY MODEL YEAR FOR LOS ANGELES

MODEL YEARS	Deterioration Factors Significantly Different from Zero		
	Pollutant	Deterioration Factor (Gm/Mi per Mile)	Significance Level
Pre-1968	HC	0.2841 -04	.90
	CO	0.3829 -03	.99
	NOXC	-.2461 -04	.995
1968, 1969	HC	0.3026 -04	.95
	CO	0.3118 -03	.75
	NOXC	-.5515 -04	.75
1970, 1971	HC	0.3772 -04	.975
	NOXC	0.7016 -04	.975
1972 - 1974	NOXC	0.5057 -04	.999

(Denver, Los Angeles, and low-altitude, non-California cities). In Table 16, it is not clear why the newer model (1972-74) exhibited a negative CO deterioration factor. Table 19 shows the significant differences between the 1972-1974 Model Year group and the earlier Model Years for the low-altitude, non-California sites. The results in general are consistent with those of Table 5 showing Model Year differences in the FTP data and are, in part, due to the lack of deterioration data for the late model vehicles.

Table 20 shows the significant deterioration differences between the test sites for the 1972-1974 Model Years. These trends do not correspond at all with the differences experienced in the FTP data for 1971 vehicles (Table 11).

4.2.2 Emission Deterioration as Reflected in the Modal Emission Model Coefficients

The modal emission model relates emission rate $\dot{e}$ to a linear combination of velocity and acceleration terms. The general form of the model is

$$\dot{e}(V,A) = h(a)\ e_{A/D} + [1-h(a)]\ \dot{e}_{ss}$$

where

$\dot{e}_{A/D} = a_1 + a_2V + a_3A + a_4VA + a_5V^2 + a_6A^2 + a_7VA^2 + a_8VA^2 + a_9V^2A^2$ for acceleration/deceleration modes;

$\dot{e}_{ss} = a_{10} + a_{11}V + a_{12}V^2$ for steady-state modes (A=0);

$h(a)$ is a linear weighting function which combines the two portions in the region close to A=0.

The a_i coefficients or weighting factors in the model have been determined for each vehicle in the FY 71 and FY 73 EFP's (References 1 and 2). The emission rate over any specified driving sequence can be estimated using these functions. Further analysis under EPA Contract No. 68-03-0435 indicates, however, that the constant term (a_1) and the acceleration term (a_3) may account for more than 80% of the variance in the data. In addition, the velocity term (a_2) may account for up to another 5%. Thus, it may be possible to provide an estimate of emission rate over a specified driving sequence using a simpler model without much loss in precision.

Table 19

SDS SIGNIFICANT DETERIORATION DIFFERENCES:
MODEL YEAR EFFECTS (LOW-ALTITUDE, NON-CALIFORNIA CITIES)

Model Year Test	Significant Differences Between Deterioration Factors	
	Pollutant Deterioration Difference	Significance Level
Pre-1968 vs. (1972-1974)	$CO_{Pre-68} > CO_{72-74}$	.75
	$NOXC_{Pre-68} < NOXC_{72-74}$	.975
(1968, 1969) vs. (1972 - 1974)	$HC_{68,\ 69} > HC_{72-74}$	.90
	$CO_{68,\ 69} > CO_{72-74}$	.998
	$NOXC_{68,69} < NOXC_{72-74}$	.999
(1970, 1971) vs. (1972 - 1974)	$HC_{70,\ 71} > HC_{72-74}$	.98
	$CO_{70,\ 71} > CO_{72-74}$	.98
	$NOXC_{70,71} < NOXC_{72-74}$	.99

Table 20

SDS SIGNIFICANT DETERIORATION DIFFERENCES:
TEST SITE EFFECTS (MODEL YEARS 1972-1974)

Test Site Test	Significant Differences Between Deterioration Factors	
	Pollutant Deterioration Difference	Significance Level
Low Altitude, Non-Calif. vs. Denver	$HC_{Low} < HC_{Denver}$ $CO_{Low} < CO_{Denver}$	.75 .75
Low Altitude, Non-Calif. vs. Los Angeles	$CO_{Low} < CO_{LA}$	.75
Denver vs. Los Angeles	$HC_{Denver} > HC_{LA}$ $CO_{Denver} > CO_{LA}$	.75 .75

The observations above indicate that emission deterioration due to age or mileage accumulation should also be reflected as changes in the important model coefficients over comparable age or mileage periods. Consequently, linear regression analysis was conducted on the a_1, a_2, and a_3 coefficients as functions of the mileage of the 1972-1974 vehicles. This analysis was done for each of the test site groups. Statistically significant changes in these model parameters with mileage are listed in Tables 21-23. To aid in relating these mileage trends in model parameters to the bag value emissions deterioration, Table 24 summarizes the comparable bag value factors and component factors. Several patterns emerge:

1. Wherever significant trends exist in both the a_1 and a_2 components, these trends are of opposite sign.

2. Wherever significant mileage effects exist for both the bag value and a_1 coefficient, these effects have the same sign in the case of HC and CO emissions, but the opposite sign in the case of NOX emissions.

3. CO deterioration does not show up as a change in the (acceleration) term; significant trends in the a_1 (constant) term, however, are exhibited for CO in each case.

4. Both bag value emission deterioration and emission model coefficient changes are less evident in the Los Angeles tested vehicles than in vehicles from other test sites.

These observations are tentative at best; their significance is not understood at this point. Under less stringent time and funding constraints, further studies of the model emissions data might yield additional relationships. Such analysis could include that half of the test vehicles with low accumulated mileage as an observation set to model emissions and predict how the vehicles with higher mileage will perform. Comparison of the predictions with the actual emission levels of the higher mileage set would indicate the feasibility of using a "mileage adaptive" modal emissions model to predict vehicle emissions over any specified driving sequence, at some future time.

Table 21

SIGNIFICANT CHANGES WITH MILEAGE IN THE MODAL EMISSION MODEL COEFFICIENT A1 (CONSTANT TERM)

TEST SITE	Coefficient Changes Significantly Different from Zero		
	Pollutant	Coefficient Change (Δa_1 per Mile)	Significance Level
Low Altitude, Non-Calif. Cities	HC	0.1859 -06	.90
	CO	-.2983 -05	.90
Denver	HC	0.9960 -06	.999
	CO	0.2393 -04	.995
	NOXC	-.3079 -06	.975
Los Angeles	CO	0.3688 -05	.75
	NOXC	-.2306 -06	.75

Table 22

SIGNIFICANT CHANGES WITH MILEAGE IN THE MODAL EMISSION MODEL COEFFICIENT A2 (VELOCITY TERM)

TEST SITE	Coefficient Changes Significantly Different from Zero		
	Pollutant	Coefficient Change (Δa_2 per Mile)	Significance Level
Low Altitude, Non-Calif. Cities	HC	-.9885 -08	.75
	CO	0.3213 -06	.975
	NOXC	0.9473 -08	.75
Denver	HC	-.3379 -07	.95
	CO	-.1028 -05	.90
	NOXC	0.2333 -07	.90
Los Angeles	None	--	--

Table 23

SIGNIFICANT CHANGES WITH MILEAGE IN THE MODAL EMISSION MODEL COEFFICIENT A3 (ACCELERATION TERM)

TEST SITE	Coefficient Changes Significantly Different from Zero		
	Pollutant	Coefficient Change (Δa_3 per Mile)	Significance Level
Low Altitude, Non-Calif. Cities	HC	0.2617 -07	.90
	NOXC	-.2150 -07	.75
Denver	NOXC	0.3929 -07	.75
Los Angles	HC	0.5672 -07	.90

Table 24

COMPARISON OF SDS BAG VALUE DETERIORATION FACTORS
WITH MODAL EMISSIONS MODEL COMPONENT TRENDS

Test Site	Pollutant	Significant Bag Value Deteriorations	Emissions Model Component Trends with Mileage		
			a_1 (constant)	a_2 (velocity)	a_3 (acceleration)
Low Altitude, Non-California Cities	HC	---	0.1859 -06	-.9885 -08	0.2617 -07
	CO	-.2411 -03	-.2983 -05	0.3213 -06	---
	NOXC	0.4119 -04	---	0.9473 -08	-.2150 -07
Denver	HC	0.4890 -04	0.9960 -06	-.3379 -07	---
	CO	0.1020 -02	0.2393 -04	-.1028 -05	---
	NOXC	0.3875 -04	-.3079 -06	0.2333 -07	0.3929 -07
Los Angeles	HC	---	---	---	0.5672 -07
	CO	---	0.3688 -05	---	---
	NOXC	0.5057 -04	-.2306 -06	---	---

4.2.3 Observed Idle Mode Bag Values -- Changes with Mileage

A separate analysis of idle mode bag values was made to compare with the total bag value emission deterioration rates. Table 25 shows those deterioration factors for each test site which were statistically different from zero. Table 26 lists the significant deterioration rate differences between test sites. These patterns are quite consistent with the differences noted in Table 20 for total observed bag value deterioration rates, except that no significant differences between the low altitude, non-California test sites and Los Angeles were found.

Table 25

SDS SIGNIFICANT DETERIORATION FACTORS IN IDLE MODE BAG VALUES

BY TEST SITE (MODEL YEARS 1972-1974)

	Deterioration Factors Significantly Different from Zero		
Test Site	Pollutant	Deterioration Factor (Gm/Mi per Mile)	Significance Level
Low Altitude, Non-Calif. Cities	None	--	--
Denver	HC	0.1002 -04	.99
	CO	0.2740 -03	.975
Los Angeles	NOXC	-.6376 -06	.75

Table 26

SDS IDLE MODE SIGNIFICANT DETERIORATION DIFFERENCES:

TEST SITE EFFECTS (MODEL YEARS 1972-1974)

Test Site Test	Significant Differences Between Deterioration Factors	
	Pollutant Deterioration Difference	Significance Level
Low Altitude, Non-Calif. vs. Denver	$HC_{Low} < HC_{Denver}$ $CO_{Low} < CO_{Denver}$	.95
Low Altitude, Non-Calif. vs. Los Angeles	No significant differences	--
Denver vs. Los Angeles	$HC_{Denver} > HC_{LA}$ $CO_{Denver} > CO_{LA}$	.95 .90

Section 5

PROBLEMS AND RECOMMENDATIONS

This section reviews the problems of predicting emissions with available EFP data including the engineering and statistical significance of the predictions and the danger of extrapolating beyond age and/or mileage ranges. Some of the problems are general to any prediction process; others are specific to the EFP design and format. In all cases the final use or application of the deterioration estimates influences the relative importance of these problems and dictates the appropriate remedies. The problems discussed below apply to both direct emission estimates and emission estimates derived by means of a modal emissions model.

5.1 USING EFP DATA TO DETERMINE HISTORICAL EMISSION TRENDS

This study has shown that it is possible to characterize samples from vehicle populations in terms of overall emission trends with mileage and age. Furthermore, these emission trends, on a population basis, can be assumed to be linear with mileage. It is reasonable therefore to use the EFP data to answer questions of the type "What is the overall emission deterioration of the entire vehicle population over the EFP test years?" or "How did 1971 model vehicles, as a group, deteriorate over the EFP test years?" The answers to these questions do not involve predicting emission levels, or extrapolating to years or mileages outside the tested ranges, but merely require analysis of historical trends. Neither do these questions require knowledge of how the emissions of individual vehicles have changed or are expected to change.

The EFP data itself does not contain deterioration information for individual vehicles. The reason that the data can be used to indicate overall emission trends can be seen by considering the EFP vehicles as random samples from, say, three subsets of vehicles -- (1) those with very low mileage, (2) those with moderate mileage, and (3) those with high mileage. The samples from each subset can be used to estimate the mean or average emission level for that subset. Assume now that the estimated mean of

subset (2) is greater than the estimated mean of subset (1). We can conclude that if the only difference in the character of the two subsets is accumulated mileage, then the difference in the means must be a result of mileage accumulation. Even though no emission measurements at the low mileages have been made for those vehicles from subset (2), we assume that the results for the low mileage vehicles in subset (1) are indicative of what those measurements would have been. Similarly, we can compare the estimated mean of subset (3) with those of subsets (1) and (2). In reality the EFP consists of samples from many "subsets" of vehicles spanning the range of accumulated mileages. Linear regression analysis is used to estimate the most consistent trend over all the subsets.

What are the problems with this procedure? First, it is not possible to assume that the only difference among the subsets is accumulated mileage. Vehicle age, for example, is also expected to be different. This problem may be insignificant for vehicles from one model year. Even in samples taken from several model years, it appears that either mileage or age can be used as the basis for emission deterioration analysis; simultaneous consideration of both parameters is not necessary (see Section 3.4).

A more significant problem arises when one wishes to determine the emission trends from a very restricted class of vehicles, for example, 1971 Chevrolets with 250-300 CID engines. The number of vehicles in the EFP representing this class would be too small to yield adequate sample sizes in the mileage accumulation "subsets." The result would be a very low level of confidence in the estimate of each subset mean and, therefore, in the overall emission trend estimates. Furthermore, without being able to adequately analyze deterioration results for small groups of vehicles, it is impossible to predict or extrapolate the emissions for particular mixtures of these groups.

Another consideration, although not specifically a problem, is the engineering significance of the deterioration estimates. For data sets where the statistical significance level is low, the deterioration factor must be accepted to be zero. This situation results when the data indeed indicates that the slope of the regression is very close to zero or when the data set is small and the variance is large. This latter case can result in non-zero slope estimates, but the estimates are too uncertain to be accepted. For some data sets, on the other hand, especially when the sample size is large, a statistically significant result may have questionable engineering importance. With enough data, even a slight

deviation from a zero slope will be detected in the regression analysis. In all cases it is necessary to consider the magnitude of the deterioration estimate in terms of its impact, in an engineering sense, on emission levels.

This point can be illustrated by considering the NOX deterioration results from two data sets. From the FY 71 EFP, the regression on 1971 Model Year vehicles is:

NOX (gm/mi) = 4.914 (gm/mi) + 0.05388 (gm/mi) per 10,000 miles

This regression shows a deterioration factor of approximately 1% per 10,000 miles. In addition to being insignificant in an engineering sense, this estimate based on 74 observations is statistically insignificant (slope F-ratio = 0.05). Next, consider the overall regression on all vehicles in the three EFPs:

NOX (gm/mi) = 4.409 (gm/mi) -0.02870 (gm/mi) per 10,000 miles

This regression shows an even smaller deterioration factor, approximately 3/4%; but because the estimate is based on 1667 observations, the results are listed as statistically significant (slope F-ratio = 2.91).

5.2 USING EFP DETERIORATION ESTIMATES TO PREDICT OVERALL EMISSION LEVELS

Predicting emission levels over future mileage or age ranges requires consideration of the expected error in extrapolating historical emission trends to future time. If one is concerned with the general quality of air at some future date, it is tempting to substitute the appropriate ages and mileages in the regression analysis equations. Prediction errors, however, can distort the estimates in several ways.

The first problem is the uncertainty with which the historical emission trend has been determined. It is a straightforward procedure to determine the confidence bounds on the regression or on the mean of the observations at a specified accumulated mileage range. As the specified mileage or age is taken to be farther and farther from the mean mileage

or age associated with the historical data, the uncertainty in the emission level estimates, as expressed by the confidence bounds, becomes larger and larger. To illustrate this concept, consider the regression of HC emissions with mileage for 1971 Model Year vehicles (three EFPs). At the mean accumulated mileage value for the historical data set (27,350 miles), the regression estimate is 3.86 gm/mi. The 95% confidence interval on the mean of observations at this mileage is at a minimum and equal to ±.23 gm/mi, or about ±5.9% of the estimate. As the specified mileage is increased to 50,000 or 100,000 miles, the confidence bound about the regression estimate, i.e. the uncertainty, grows. Table 27 shows that at 100,000 miles the 95% confidence interval is ±19.6% of the regression estimate. For CO emissions the results are essentially the same, but for NOX emissions the confidence interval is greater than ±30% of the estimate at 100,000 miles.

Similar results are obtained for emissions estimated using age regressions.

The second problem is a most important one and arises in the estimation of the vehicle mix at some future time. The number of vehicles of specific types and model years must be estimated, including vehicles (and estimates of their associated emission levels) which do not exist now but will be on the road in the future. Each of these estimates involving population dynamics will contain uncertainties which further compound the prediction errors.

Even if a reasonable estimate of the vehicle mix can be made, it is still necessary to estimate the accumulated mileage levels which will exist for each of the vehicle groups. This problem of course does not exist if age is used as the emission deterioration independent variable. All vehicles of a specific model year will be at age ≈ (test year - model year). The model year mix must be estimated, however, just as with the mileage regression estimates.

A final problem exists in that future emission deterioration rates may change over future intervals. Even though deterioration with mileage appears to be linear for the model years and EFPs considered in this study, there is no way to guarantee that these rates will not change or that the deterioration rates of newer vehicles with different emissions control equipment will deteriorate linearly.

Table 27

PREDICTION ERROR FOR 1971 MODEL YEAR EMISSIONS

FTP Data - EFP FY 71-73

Low-Altitude, Non-California Cities

	Mileage x^*	Estimate from Regression $\bar{y} = a + bx^*$	.95 Confidence on $\bar{y}$ at x^* Interval	Relative
HC	27,350 = $\bar{x}$	3.86 gm/mi	± .23 gm/mi	± 5.9%
	50,000	4.55	± .45	± 9.9
	75,000	5.30	± .79	±14.9
	100,000	6.06	±1.19	±19.6
CO	27,350 = $\bar{x}$	52.65 gm/mi	± 3.5 gm/mi	± 6.6%
	50,000	64.28	± 6.5	±10.1
	75,000	77.13	±12.2	±15.8
	100,000	89.97	±18.1	±20.1
NOX	27,350 = $\bar{x}$	4.61 gm/mi	± .26 gm/mi	± 5.6%
	50,000	4.53	± .50	±11.0
	75,000	4.45	± .95	±21.3
	100,000	4.36	± 1.42	±32.6

5.3 USING EFP DETERIORATION ESTIMATES TO PREDICT EMISSION LEVELS OF INDIVIDUAL VEHICLES

For any analysis which requires an estimate of emission changes for individual vehicles, the EFP data in its current format is not applicable. This limitation prevents addressing questions of the form "What percentage of 1971 vehicles are deteriorating at a rate twice as high as the average?" or "What extremes of deterioration are experienced in catalyst equipped vehicles?" or "How do the deterioration rates of vehicles which had emission levels below the Federal Standards at 10,000 miles compare to the deterioration rates of vehicles which exceeded the Standards in at least one pollutant" or "What is the correlation between HC, CO, and NOx deterioration rates?"

These and other questions can be answered only by first generating deterioration estimates for individual vehicles. The vehicles must be tested over a period of time at least twice, preferably three or four times. The deterioration rates for a group of vehicles then can be presented as a frequency distribution of deterioration values. With appropriate analysis, the mean rate for the group can be determined along with other statistics necessary to address questions of the type listed above. The disadvantage of this procedure of course is the increased cost of testing and recalling the vehicles.

5.4 RECOMMENDATIONS FOR FUTURE STUDIES

The primary concern for future studies should be the identification of the expected or the required goals of using the EFP data in establishing deterioration rates. As discussed above, if only historical trends are of interest, the current EFP design can be used. Additional attention is necessary, however, in specifying the types and numbers of vehicles which should be tested in specific groups so as to produce statistically significant results where required.

If, however, a prediction capability is desired, either in terms of specific vehicle groups or in terms of the entire vehicle population, further studies are recommended:

1. Consideration of historical, current, and postulated vehicle mixes in terms of model years, engine types, emission control equipment, and expected accumulated mileage levels per year.

2. Consideration of the number of vehicles from a specific group which are required to be tested over a period of time to yield emission predictions of specified precision, and consideration of how these should be stratified with respect to the current and future populations.

3. Consideration of the form of the emission deterioration curves for newer vehicles.

4. Comparison of emission level predictions derived from the FY 71-73 EFP data as contained in this study with actual emission levels measured at this time (1976).

If, in fact, it is desired to determine the factors which contribute to emission deterioration or to design and evaluate inspection/maintenance procedures in order to reduce deterioration, additional tests or a redesign of the EFP test format is required. Emission deterioration rates must be determined for individual vehicles and the maintenance of these vehicles must be closely controlled. Limited data from such a test program exists and has been reported (Reference 8). It is necessary to trade off the costs of similar testing with the expected gains in information.

An additional advantage in measuring emissions deterioration for individual vehicles is that it becomes possible to analyze the correlation between the deteriorations of the three emittants. Just as there exists a correlation between emission levels at any one time, it is suspected that the rates of change in these levels over a period of time are also correlated. It may be possible to develop a theory of composite deterioration by applying Principal Component Analysis and Factor Analysis procedures to the existing data reported in Reference 8. If the results are promising, additional testing and studies should be planned.

REFERENCES

1. Paul Kunselman, H.T. McAdams, C.J. Domke, and Marcia Williams, Automobile Exhaust Emmission Modal Analysis Model, Environmental Protection Agency Report. No. EPA-460/3-74-005 (January 1974).

2. H.T. McAdams, Automobile Exhaust Emission Modal Analysis Model Extension and Refinement, Environmental Protection Agency Report No. EPA-460/3-74-024 (October 1974).

3. Automobile Exhaust Emission Surveillance--A Summary, Environmental Protection Agency Report No. APTD-1544 (March 1973).

4. Marcia E. Williams, John T. White, Lois A. Platte, Charles J. Domke, Automobile Exhaust Emission Surveillance--Analysis of the FY 72 Program, Environmental Protection Agency Report No. EPA-460/2-74-001 (February 1974).

5. Jeffrey Bernard, Paul Donovan, H.T. McAdams, Automobile Exhaust Emission Surveillance--Analysis of the FY 73 Program, Environmental Protection Agency Report No. EPA-460/3-75-007 (July 1975).

6. Norman Draper and Harry Smith, Applied Regression Analysis, John Wiley and Sons, Inc. (1966).

7. K. Diem and C. Lenther, Editors, Scientific Tables, GEIGY Pharmaceuticals Division of CIBA-GEIGY Corp. (1970).

8. J.A. Gunderson and L. Resnick, Degradation Effects on Motor Vehicle Exhaust Emission, Society of Automotive Engineers Congress and Exposition Report No. 760366 (1976).

APPENDIX I

T-TEST FOR THE EQUALITY OF TWO SLOPES

(Adapted from Reference 7)

Given two regression:

$Y_1 = a_1 + b_1x_1$ $\qquad$ n_1 = no of data points

$Y_2 = a_2 + b_2x_2$ $\qquad$ n_2 = no of data points

Test the hypothesis that $b_1 = b_2$, under the conditions that the residual variances $(\sigma^2_{y \cdot x})_1$ and $(\sigma^2_{y \cdot x})_2$ are unknown and unequal.

The test statistic is given as $\dfrac{b_1 - b_2}{S_{D_b}}$

and is compared to a t variable with ν_t degress of freedom, where

$$S_{D_b} = \frac{(S^2_{y \cdot x})_1}{S_{x_1}} + \frac{(S^2_{y \cdot x})_2}{S_{x_2}}$$

$$S_{x_1} = \text{var}(x_1) \cdot (n_1 - 1)$$

$$S_{x_2} = \text{var}(x_2) \cdot (n_2 - 1)$$

$(S^2_{y \cdot x})_1$ = estimate of $(\sigma_{y \cdot x})_1$ = residual mean square (1)

$(S^2_{y \cdot x})_2$ = estimate of $(\sigma_{y \cdot x})_2$ = residual mean square (2)

$$\nu_t = \frac{1}{K^2/\nu_1 + (1-K)^2/\nu_2}$$

$$\nu_1 = n_1 - 2$$

$$\nu_2 = n_2 - 2$$

$$K = \frac{(S^2_{y \cdot x})_1 \cdot S_{x_2}}{(S^2_{y \cdot x})_1 \cdot S_{x_2} + (S^2_{y \cdot x})_2 \cdot S_{x_1}}$$

APPENDIX II

LINEAR REGRESSION RESULTS FOR FTP DATA

The tables in this appendix list the results of linear regression analysis on FTP data from the FY 71-73 Emission Factor Programs. Regression results are given for both the basic data sets and corrected data sets with 10% outliers removed. The intercept and slope data is presented in scientific notation with the power of ten listed after the four-place characteristic. Calculated F-ratios for testing the significance of the slope (deterioration factor) are also included.

Table II-1

REGRESSION RESULTS BY MODEL YEAR

EFP FY 71-73

Low-Altitude, Non-California Cities

MODEL YEAR	DATA	POLLUTANT	INTERCEPT	SLOPE	SLOPE F-RATIO
Pre-1968	BASIC (N=329)	HC	0.7800 +01	0.2306 -04	2.13
		CO	0.8497 +02	0.3590 -03	8.59
		NOXC	0.3365 +01	0.1096 -05	0.07
	CORRECTED (N=301)	HC	0.7334 +01	0.1486 -04	2.79
		CO	0.8761 +02	0.2754 -03	5.91
		NOXC	0.3484 +01	-.1990 -05	0.26
1970	BASIC (N=266)	HC	0.4247 +01	0.2884 -04	4.25
		CO	0.5303 +02	0.3717 -03	7.79
		NOXC	0.4824 +01	-.5786 -05	0.85
	CORRECTED (N=242)	HC	0.3790 +01	0.2267 -04	13.21
		CO	0.4636 +02	0.4332 -03	17.47
		NOXC	0.4843 +01	-.8343 -05	2.41
1971	BASIC (N=308)	HC	0.3347 +01	0.3508 -04	12.87
		CO	0.4405 +02	0.4821 -03	11.62
		NOXC	0.4674 +01	-.3170 -05	0.20
	CORRECTED (N=285)	HC	0.3038 +01	0.3018 -04	30.32
		CO	0.3860 +02	0.5137 -03	23.88
		NOXC	0.4703 +01	-.3384 -05	0.25

TABLE II-1
(continued)

MODEL YEAR	DATA	POLLUTANT	INTERCEPT	SLOPE	SLOPE F-RATIO
1972 (FY 72 & 73)	BASIC (N=260)	HC	0.3366 +01	0.3003 -04	3.35
		CO	0.4171 +02	0.5710 -03	10.35
		NOXC	0.4375 +01	0.1181 -04	2.01
	CORRECTED (N=246)	HC	0.3074 +01	0.2536 -04	15.56
		CO	0.4003 +02	0.4710 -03	12.22
		NOXC	0.4515 +01	0.9674 -06	0.02
1973 (FY 73)	BASIC (N=140)	HC	0.4098 +01	-.1347 -05	0.01
		CO	0.5571 +02	-.1277 -03	0.24
		NOXC	0.3215 +01	0.1399 -04	1.27
	CORRECTED (N=129)	HC	0.3692 +01	0.3876 -05	0.16
		CO	0.4862 +02	-.1935 -04	0.01
		NOXC	0.3162 +01	0.9978 -05	1.28

Table II-2

REGRESSION RESULTS BY FISCAL YEAR

Model Year 1971

Low-Altitude, Non-California Cities

EFP FY	DATA	POLLUTANT	INTERCEPT	SLOPE	SLOPE F-RATIO
1971	BASIC (N=80)	HC	0.2956 +01	0.3004 -04	2.25
		CO	0.4562 +02	0.4597 -04	0.01
		NOXC	0.5074 +01	-.5312 -05	0.05
	CORRECTED (N=74)	HC	0.2907 +01	0.1539 -04	1.10
		CO	0.4039 +02	0.1675 -03	0.26
		NOXC	0.4914 +01	0.5388 -05	0.05
1972	BASIC (N=120)	HC	0.4241 +01	0.6382 -05	0.07
		CO	0.6184 +02	-.5377 -04	0.02
		NOXC	0.3961 +01	0.1049 -04	0.49
	CORRECTED (N=112)	HC	0.3574 +01	0.1377 -04	1.23
		CO	0.5614 +02	-.1092 -04	0.00
		NOXC	0.4038 +01	0.1286 -04	0.71
1973	BASIC (N=108)	HC	0.3997 +01	0.2291 -04	1.51
		CO	0.3771 +02	0.6471 -03	8.50
		NOXC	0.4736 +01	-.1654 -05	0.02
	CORRECTED (N=97)	HC	0.3396 +01	0.2280 -04	5.77
		CO	0.2631 +02	0.8169 -03	25.07
		NOXC	0.4778 +01	-.3604 -05	0.10

Table II-3

REGRESSION RESULTS BY FISCAL YEAR

All Model Years

Low-Altitude, Non-California Cities

EFP FY	DATA	POLLUTANT	INTERCEPT	SLOPE	SLOPE F-RATIO
1971	BASIC (N=412)	HC	0.2546 +01	0.8737 -04	62.41
		CO	0.3928 +02	0.8683 -03	79.56
		NOXC	0.5283 +01	-.2052 -04	24.11
	CORRECTED (N=377)	HC	0.2701 +01	0.6542 -04	127.99
		CO	0.3732 +02	0.7634 -03	101.61
		NOXC	0.5285 +01	-.2066 -04	27.47
1972	BASIC (N=680)	HC	0.3023 +01	0.6935 -04	91.15
		CO	0.4214 +02	0.7033 -03	124.40
		NOXC	0.4563 +01	-.1002 -04	15.39
	CORRECTED (N=630)	HC	0.2645 +01	0.5899 -04	296.88
		CO	0.3846 +02	0.6868 -03	187.04
		NOXC	0.4666 +01	-.1183 -04	23.90
1973	BASIC (N=720)	HC	0.3616 +01	0.5429 -04	51.03
		CO	0.4606 +02	0.6660 -03	102.40
		NOXC	0.3841 +01	0.1425 -04	21.82
	CORRECTED (N=658)	HC	0.3598 +01	0.3585 -04	97.46
		CO	0.4278 +02	0.5909 -03	136.55
		NOXC	0.3712 +01	0.1542 -04	32.79

Table II-4

REGRESSION RESULTS FOR ALL MODEL YEARS
FOR EFP FY 71-73

DATA	POLLUTANT	INTERCEPT	SLOPE	SLOPE F-RATIO
BASIC (N=1812)	HC	0.3190 +01	0.6657 -04	196.28
	CO	0.4341 +02	0.7170 -03	301.17
	NOXC	0.4414 +01	-.2304 -05	1.58
CORRECTED (N=1667)	HC	0.3006 +01	0.5243 -04	499.71
	CO	0.3988 +02	0.6678 -03	424.88
	NOXC	0.4409 +01	-.2870 -05	2.91

Table II-5

REGRESSION RESULTS BY TEST SITE

EFP FY 71-73

CITY	DATA	POLLUTANT	INTERCEPT	SLOPE	SLOPE F-RATIO
LOW ALTITUDE, NON-CALIF.	BASIC (N=308)	HC	0.3347 +01	0.3508 -04	12.87
		CO	0.4405 +02	0.4821 -03	11.62
		NOXC	0.4674 +01	-.3170 -05	0.20
	CORRECTED (N=285)	HC	0.3038 +01	0.3018 -04	30.32
		CO	0.3860 +02	0.5137 -03	23.88
		NOXC	0.4703 +01	-.3384 -05	0.25
DENVER	BASIC (N=77)	HC	0.6074 +01	0.5763 -05	0.12
		CO	0.1014 +03	-.9555 -04	0.09
		NOXC	0.3010 +01	-.2751 -05	0.06
	CORRECTED (N=70)	HC	0.5634 +01	0.1427 -04	1.11
		CO	0.9775 +02	-.8694 -04	0.10
		NOXC	0.2901 +01	-.2144 -05	0.06
LOS ANGELES	BASIC (N=78)	HC	0.3497 +01	0.2146 -04	2.83
		CO	0.6079 +02	0.5459 -04	0.05
		NOXC	0.3927 +01	-.1092 -04	2.38
	CORRECTED (N=69)	HC	0.3188 +01	0.1958 -04	5.58
		CO	0.5069 +02	0.2180 -03	1.74
		NOXC	0.4101 +01	-.1411 -04	4.48

Table II-6

REGRESSION RESULTS BY CID

EFP FY 71-73

Low-Altitude, Non-California Cities

CID RANGE	DATA	POLLUTANT	INTERCEPT	SLOPE	SLOPE F-RATIO
0-249	BASIC (N=90)	HC	0.2079 +01	0.7691 -04	10.54
		CO	0.4298 +02	0.4238 -04	0.03
		NOXC	0.4485 +01	-.2209 -04	1.79
	CORRECTED (N=85)	HC	0.2518 +01	0.4312 -04	8.73
		CO	0.3931 +02	0.1305 -03	0.43
		NOXC	0.4905 +01	-.4295 -04	7.37
250-360	BASIC (N=134)	HC	0.3924 +01	0.1896 -04	1.93
		CO	0.3800 +02	0.6793 -03	19.91
		NOXC	0.4837 +01	-.7817 -05	0.93
	CORRECTED (N=121)	HC	0.3293 +01	0.2533 -04	12.78
		CO	0.3553 +02	0.6518 -03	25.56
		NOXC	0.4717 +01	-.4506 -05	0.33
361-500	BASIC (N=84)	HC	0.3643 +01	0.2751 -04	2.70
		CO	0.6784 +02	0.1305 -03	0.13
		NOXC	0.5187 +01	0.1234 -05	0.01
	CORRECTED (N=77)	HC	0.3614 +01	0.1510 -04	2.04
		CO	0.6368 +02	0.5027 -04	0.03
		NOXC	0.5284 +01	-.5367 -06	0.00

Table II-7

REGRESSION RESULTS BY ACCUMULATED MILEAGE RANGE

EFP FY 71-73

Low-Altitude, Non-California Cities

MILEAGE ACCUMULATION RANGE (Miles per year)	DATA	POLLUTANT	INTERCEPT	SLOPE	SLOPE F-RATIO
0-9,999	BASIC (N=105)	HC	0.3732 +01	0.2522 -04	0.84
		CO	0.5078 +02	0.3552 -03	0.66
		NOXC	0.5034 +01	-.3620 -04	2.98
	CORRECTED (N=97)	HC	0.3039 +01	0.3768 -04	5.75
		CO	0.5036 +02	0.9265 -04	0.08
		NOXC	0.4910 +01	-.2751 -04	1.66
10,000-13,999	BASIC (N=104)	HC	0.2024 +01	0.8083 -04	12.80
		CO	0.2735 +02	0.8667 -03	17.47
		NOXC	0.4589 +01	0.1109 -04	0.53
	CORRECTED (N=97)	HC	0.2403 +01	0.5100 -04	23.52
		CO	0.2924 +02	0.7030 -03	15.35
		NOXC	0.4483 +01	0.1754 -04	1.34
14,000-100,000	BASIC (N=99)	HC	0.3519 +01	0.2411 -04	3.24
		CO	0.4698 +02	0.4348 -03	2.84
		NOXC	0.4925 +01	-.1204 -04	1.44
	CORRECTED (N=93)	HC	0.3258 +01	0.2051 -04	5.07
		CO	0.3741 +02	0.5820 -03	9.85
		NOXC	0.5067 +01	-.1340 -04	1.75

Table II-8

AGE AND MILEAGE REGRESSION RESULTS

EFP FY 71-73

Low-Altitude, Non-California Cities

INDEPENDENT VARIABLE	DATA	POLLUTANT	INTERCEPT	SLOPE	SLOPE F-RATIO
MILEAGE	BASIC (N=308)	HC	0.3347 +01	0.3508 -04	12.87
		CO	0.4405 +02	0.4821 -03	11.62
		NOXC	0.4674 +01	-.3170 -05	0.20
	CORRECTED (N=285)	HC	0.3038 +01	0.3018 -04	30.32
		CO	0.3860 +02	0.5137 -03	23.88
		NOXC	0.4703 +01	-.3384 -05	0.25
AGE	BASIC (N=308)	HC	0.3069 +01	0.4474 -01	15.12
		CO	0.4421 +02	0.4716 +00	7.87
		NOXC	0.4794 +01	-.7463 -02	0.80
	CORRECTED (N=286)	HC	0.3085 +01	0.2851 -01	18.43
		CO	0.4282 +02	0.3568 +00	7.80
		NOXC	0.4854 +01	-.8365 -02	1.04

APPENDIX III

LINEAR REGRESSION RESULTS FOR SDS BAG VALUE DATA AND FOR SELECTED MODAL EMISSION MODEL PARAMETERS

The tables in this appendix list the results of linear regression analysis on SDS data from the FY 71 and FY 73 Emission Factor Programs. Regression results are given for both the basic data sets and corrected data sets with 10% outliers removed. The intercept and slope data is presented in scientific notation with the power of ten listed after the four-place characteristic. Calculated F-ratios for testing the significance of the slopes (deterioration factors) are also included.

The first four tables are observed bag value emissions vs. mileage regressions. The next three tables are mileage regressions on the a_1, a_2, and a_3 coefficients from the modal emissions model for Model Year 72-74 vehicles (low-altitude, non-California test sites). The last table gives mileage regressions for observed idle mode bag values.

Table III-1

BAG VALUE REGRESSION RESULTS FOR PRE-1968 VEHICLES

CITY	DATA	POLLUTANT	INTERCEPT	SLOPE	SLOPE F-RATIO
LOW ALTITUDE, NON-CALIF.	BASIC (N=385)	HC	0.4825 +01	0.2452 -04	2.94
		CO	0.6051 +02	0.5628 -04	0.51
		NOXC	0.5043 +01	-.5229 -05	1.03
	CORRECTED (N=351)	HC	0.5089 +01	0.2831 -05	0.24
		CO	0.5940 +02	0.1028 -04	0.02
		NOXC	0.5212 +01	-.7901 -05	2.91
DENVER	BASIC (N=97)	HC	0.7095 +01	0.1383 -04	0.44
		CO	0.1285 +03	0.9289 -06	0.00
		NOXC	0.1800 +01	0.7648 -05	1.01
	CORRECTED (N=91)	HC	0.7586 +01	-.7270 -06	0.00
		CO	0.1201 +03	0.9025 -04	0.14
		NOXC	0.1782 +01	0.6088 -05	1.01
LOS ANGELES	BASIC (N=106)	HC	0.2967 +01	0.5060 -04	2.55
		CO	0.3296 +02	0.3878 -03	5.22
		NOXC	0.5882 +01	-.2547 -04	8.93
	CORRECTED (N=95)	HC	0.3427 +01	0.2841 -04	3.86
		CO	0.3085 +02	0.3829 -03	7.77
		NOXC	0.5945 +01	-.2461 -04	9.21

Table III-2

BAG VALUE REGRESSION RESULTS FOR MODEL YEARS 1968 & 1969

CITY	DATA	POLLUTANT	INTERCEPT	SLOPE	SLOPE F-RATIO
LOW ALTITUDE, NON-CALIF.	BASIC (N=141)	HC	0.2776 +01	0.3028 -04	2.77
		CO	0.2955 +02	0.3112 -03	2.67
		NOXC	0.6934 +01	-.1203 -04	0.89
	CORRECTED (N=130)	HC	0.2702 +01	0.1621 -04	4.72
		CO	0.2532 +02	0.2632 -03	5.04
		NOXC	0.6954 +01	-.9421 -05	0.57
DENVER	BASIC (N=35)	HC	0.5384 +01	0.2962 -05	0.02
		CO	0.9573 +02	0.2307 -03	0.18
		NOXC	0.2526 +01	0.4509 -05	0.08
	CORRECTED (N=32)	HC	0.4969 +01	0.1538 -04	0.61
		CO	0.9938 +02	0.2604 -03	0.21
		NOXC	0.1890 +01	0.1888 -04	1.60
LOS ANGELES	BASIC (N=32)	HC	0.2069 +01	0.3124 -04	4.44
		CO	0.1705 +02	0.4619 -03	3.98
		NOXC	0.8646 +01	-.5734 -04	2.74
	CORRECTED (N=30)	HC	0.1981 +01	0.3026 -04	5.26
		CO	0.2021 +02	0.3118 -03	2.35
		NOXC	0.8718 +01	-.5515 -04	2.39

Table III-3

BAG VALUE REGRESSION RESULTS FOR MODEL YEARS 1970 & 1971

CITY	DATA	POLLUTANT	INTERCEPT	SLOPE	SLOPE F-RATIO
LOW ALTITUDE, NON-CALIF.	BASIC (N=150)	HC	0.2262 +01	0.2657 -04	8.88
		CO	0.3846 +02	-.1484 -03	0.77
		NOXC	0.6448 +01	0.1104 -04	0.54
	CORRECTED (N=137)	HC	0.2177 +01	0.1974 -04	9.54
		CO	0.2721 +02	0.1679 -03	2.13
		NOXC	0.6914 +01	-.5011 -05	0.11
DENVER	BASIC (N=37)	HC	0.4905 +01	0.2338 -04	0.34
		CO	0.1309 +03	-.7526 -03	1.28
		NOXC	0.3115 +01	-.2963 -05	0.03
	CORRECTED (N=32)	HC	0.5325 +01	-.2386 -04	1.08
		CO	0.1246 +03	-.8214 -03	3.22
		NOXC	0.3094 +01	-.7741 -05	0.32
LOS ANGELES	BASIC (N=37)	HC	0.2205 +01	0.3717 -04	1.60
		CO	0.2206 +02	0.2182 -03	1.08
		NOXC	0.4868 +01	0.7000 -04	5.29
	CORRECTED (N=34)	HC	0.1920 +01	0.3772 -04	7.06
		CO	0.2456 +02	0.8014 -04	0.17
		NOXC	0.5156 +01	0.7016 -04	6.34

Table III-4

BAG VALUE REGRESSION RESULTS FOR MODEL YEARS 1972 - 1974

CITY	DATA	POLLUTANT	INTERCEPT	SLOPE	SLOPE F-RATIO
LOW ALTITUDE, NON-CALIF.	BASIC (N=300)	HC	0.3180 +01	0.4161 -05	0.17
		CO	0.5052 +02	-.1908 -03	1.76
		NOXC	0.4213 +01	0.4362 -04	21.96
	CORRECTED (N=280)	HC	0.2928 +01	0.1381 -05	0.09
		CO	0.4755 +02	-.2411 -03	4.97
		NOXC	0.4142 +01	0.4119 -04	25.70
DENVER	BASIC (N=75)	HC	0.4042 +01	0.4839 -04	6.36
		CO	0.1134 +03	0.9526 -03	1.89
		NOXC	0.1950 +01	0.3970 -04	8.68
	CORRECTED (N=69)	HC	0.3854 +01	0.4890 -04	7.77
		CO	0.1076 +03	0.1020 -02	2.54
		NOXC	0.1944 +01	0.3875 -04	10.80
LOS ANGELES	BASIC (N=75)	HC	0.3046 +01	-.3449 -05	0.02
		CO	0.2572 +02	0.2513 -03	2.81
		NOXC	0.3489 +01	0.4441 -04	7.20
	CORRECTED (N=69)	HC	0.2655 +01	-.1787 -05	0.02
		CO	0.2747 +02	0.1149 -03	0.68
		NOXC	0.3290 +01	0.5057 -04	11.73

Table III-5

REGRESSION RESULTS FOR MODAL EMISSION MODEL COEFFICIENT A1

CITY	DATA	POLLUTANT	INTERCEPT	SLOPE	SLOPE F-RATIO
LOW ALTITUDE, NON-CALIF.	BASIC (N=300)	HC	0.3149 -01	0.9707 -07	0.31
		CO	0.6007 +00	-.2220 -05	1.25
		NOXC	0.1425 -01	-.5678 -07	0.24
	CORRECTED (N=281)	HC	0.2628 -01	0.1859 -06	3.28
		CO	0.5920 +00	-.2983 -05	3.15
		NOXC	0.1427 -01	-.1358 -07	0.02
DENVER	BASIC (N=75)	HC	0.2905 -01	0.7953 -06	8.07
		CO	0.8646 +00	0.2150 -04	7.54
		NOXC	0.7282 -02	-.2634 -06	3.54
	CORRECTED (N=66)	HC	0.2254 -01	0.9960 -06	14.13
		CO	0.7308 +00	0.2393 -04	11.51
		NOXC	0.6679 -02	-.3079 -06	6.36
LOS ANGELES	BASIC (N=75)	HC	0.3501 -01	-.3233 -06	1.06
		CO	0.1671 +00	0.3000 -05	0.94
		NOXC	0.5309 -02	-.2294 -06	1.32
	CORRECTED (N=70)	HC	0.2656 -01	-.1266 -06	0.24
		CO	0.1160 +00	0.3688 -05	1.71
		NOXC	0.3114 -02	-.2306 -06	1.92

Table III-6

REGRESSION RESULTS FOR MODAL EMISSION MODEL COEFFICIENT A2

CITY	DATA	POLLUTANT	INTERCEPT	SLOPE	SLOPE F-RATIO
LOW ALTITUDE, NON-CALIF.	BASIC (N=300)	HC	-.1445 -02	-.5080 -08	0.19
		CO	-.4502 -01	0.3263 -06	4.17
		NOXC	-.9549 -03	0.1115 -07	1.35
	CORRECTED (N=275)	HC	-.1112 -02	-.9885 -08	2.07
		CO	-.4092 -01	0.3213 -06	6.66
		NOXC	-.1015 -02	0.9473 -08	1.47
DENVER	BASIC (N=75)	HC	-.1404 -02	-.4717 -07	6.89
		CO	-.5604 -01	-.1798 -05	9.07
		NOXC	-.4412 -03	0.3400 -07	6.03
	CORRECTED (N=66)	HC	-.1319 -02	-.3379 -07	4.09
		CO	-.6178 -01	-.1028 -05	3.41
		NOXC	-.3095 -03	0.2333 -07	2.82
LOS ANGELES	BASIC (N=75)	HC	-.2170 -02	0.2689 -07	1.77
		CO	-.8897 -02	-.1513 -06	0.50
		NOXC	-.3835 -03	0.1115 -07	0.41
	CORRECTED (N=69)	HC	-.1556 -02	0.8131 -08	0.20
		CO	-.1101 -01	-.3487 -07	0.03
		NOXC	0.1306 -03	-.4740 -08	0.10

Table III-7

REGRESSION RESULTS FOR MODAL EMISSION MODEL COEFFICIENT A3

CITY	DATA	POLLUTANT	INTERCEPT	SLOPE	SLOPE F-RATIO
LOW ALTITUDE, NON-CALIF.	BASIC (N=300)	HC	0.1021 -02	0.1789 -08	0.01
		CO	0.2629 -01	-.3236 -07	0.01
		NOXC	0.1571 -03	-.3852 -07	3.49
	CORRECTED (N=268)	HC	0.2330 -03	0.2617 -07	3.80
		CO	0.1204 -01	0.1918 -06	0.36
		NOXC	-.4207 -03	-.2150 -07	1.72
DENVER	BASIC (N=75)	HC	-.1206 -02	0.1465 -07	0.12
		CO	0.4494 -01	-.4502 -06	0.11
		NOXC	-.9798 -03	0.3173 -07	1.22
	CORRECTED (N=67)	HC	-.6080 -03	0.1566 -07	0.28
		CO	0.2237 -01	-.1576 -06	0.03
		NOXC	-.7142 -03	0.3929 -07	2.73
LOS ANGELES	BASIC (N=75)	HC	0.3065 -02	0.2406 -07	0.26
		CO	0.8763 -01	0.3015 -06	0.13
		NOXC	0.3785 -03	-.7023 -07	1.51
	CORRECTED (N=69)	HC	0.1565 -02	0.5672 -07	2.78
		CO	0.7234 -01	0.5770 -06	0.60
		NOXC	-.5866 -03	-.4512 -07	0.86

Table III-8

REGRESSION RESULTS FOR IDLE MODE BAG VALUES

EFP FY 73

MODEL YEARS 1972-1974

CITY	DATA	POLLUTANT	INTERCEPT		SLOPE		SLOPE F-RATIO
LOW ALTITUDE, NON-CALIF.	BASIC (N = 300)	HC	0.8391	+00	0.1521	-05	0.03
		CO	0.1274	+02	0.4092	-04	0.48
		NOXC	0.2325	+00	-.8385	-06	0.04
	CORRECTED (N = 286)	HC	0.6940	+00	0.6561	-06	0.07
		CO	0.1108	+02	0.2505	-04	0.29
		NOXC	0.1531	+00	-.3274	-06	0.09
DENVER	BASIC (N = 75)	HC	0.7163	+00	0.6781	-05	1.39
		CO	0.1122	+02	0.2895	-03	3.63
		NOXC	0.7643	-01	0.3251	-06	0.11
	CORRECTED (N = 65)	HC	0.4855	+00	0.1002	-04	8.20
		CO	0.7831	+01	0.2740	-03	6.53
		NOXC	0.6884	-01	-.1365	-06	0.07
LOS ANGELES	BASIC (N = 75)	HC	0.6276	+00	-.2769	-05	0.17
		CO	0.6332	+01	0.5689	-04	0.49
		NOXC	0.8787	-01	-.8795	-06	3.08
	CORRECTED (N = 68)	HC	0.4036	+00	0.9529	-06	0.16
		CO	0.5870	+01	0.3380	-04	0.28
		NOXC	0.7650	-01	-.6376	-06	2.60