Summary of Peer Review Comments:
In-Use Exhaust Deterioration of 1981-1993 Model Year Cars and Trucks for Use in MOBILE6

March 22, 1999

U.S. EPA Assessment and Modeling Division

Introduction

Four papers that document the in-use exhaust deterioration of 1981-1993 model year (Tier 0) cars and trucks were peer reviewed:

- M6.STE.002, "The Determination of Hot Running Emissions from FTP Bag Emissions"
- M6.STE.003, "Determination of Start Emissions as a Function of Mileage and Soak Time for 1981-1993 Model Year Light-Duty Vehicles"
- M6.EXH.001, "Determination of Running Emissions as a Function of Mileage for 1981-1993 Model Year Light-Duty Cars and Trucks"
- M6.EXH.002, "Analysis of Emissions Deterioration Using Ohio and Wisconsin IM240 Data"

Peer review comments were solicited and received from Mr. Kevin Green, U.S. Department of Transportation; Dr. Stephen L. Gerritson, Consultant; and Dr. Larry Caretto, California State University, Northridge. Even though the affiliations of Mr. Kevin Green and Dr. Larry Caretto are provided, it should be noted that these reviewers acted as independent consultants for the purpose of this review. As such, the opinions expressed do not represent those of the affiliated organizations. EPA is also awaiting comments from Dr. Joseph Norbeck, University of California, Riverside.

These peer reviewers were chosen because they have the necessary expertise by virtue of previous experience to be able to provide an informed review. They also have had no previous involvement in the development of these papers, and no potential or perceived conflict of interest. Kevin Green was also specifically recommended by the Northeast States for Coordinated Air Use Management (NESCAUM).

EPA asked that the reviewers address the following areas:

1) report clarity,
2) overall methodology,
3) appropriateness of the data sets selected,
4) data analyses conducted, including the statistical approaches used and models selected,
5) appropriateness of the conclusions, and
6) recommendations for any alternate data sets and/or analyses.

The peer reviewers were also asked to address a list of nine issues representing major choices and assumptions in the analysis:

1) model year groupings and segregation of fuel system types (applies to all papers),
2) separation of FTP measurements into start and running emissions based on results from 76 cars (addressed in M6.STE.002),
3) running emissions versus mileage follows a piecewise linear function (addressed in M6.EXH.001),
4) FTP data sets are not considered to represent the fleet, and results based on these data sets are adjusted to match Dayton, Ohio observations (addressed in M6.EXH.001),

5) Cars tested with IM240 in Dayton are considered sufficiently representative of the national fleet, within each model year and technology grouping (addressed in M6.EXH.002),
6) odometer readings on individual Dayton vehicles are considered unreliable, and estimated mean odometer values are used instead (addressed in M6.EXH.002),
7) fast pass IM240 measurements in Dayton are converted to estimates of full IM240 (addressed in M6.EXH.002),
8) estimates of full IM240 emissions are converted to estimates of running emissions (addressed in M6.EXH.002), and
9) the weightings for the fraction of high emitters (based on the cars tested with IM240 in Dayton) are applied to the start emission estimates (addressed in M6.STE.003).

This document provides a summary of the peer review comments for each paper, arranged by area and issue, followed by general comments on the set of four reports.

These independent peer review comments are being made available to assist public and stakeholder reviewers. The independent review comments and others received will be addressed together following the general comment period. A followup document containing EPA's response to all the comments will then be provided.

Peer Review Comments

M6.STE.002, "The Determination of Hot Running Emissions from FTP Bag Emissions"
[Includes responses to issue #2]
Report Clarity

a) Should include a discussion of the options available and the rationale for selecting the sample and methodology that was used. (Gerritson)
b) Introduction should address the relationship between the HR505 and real-world hot running conditions. (Green)
c) Writing style clear and easily understood; overall clarity generally good. (Gerritson, Caretto)
d) It would be helpful to include the HR505 data in Table 1. (Green, Caretto)
e) No reference for "accompanying document" mentioned on page 2. (Caretto)
f) Provide a reference for the logarithm transformation constant. Specify in Table 3a whether or not the value of D in Table 3a contains this constant. (Caretto)
g) Delete the appendices and replace with a table in the body of the report showing the characteristics of the 77 vehicles used in the regression. (Caretto)
h) Paper should contain a clear statement of the conclusions. (Gerritson)

Overall Methodology

a) Overall study design and methodology is appropriate. (Gerritson)
b) Specify whether vehicles 221 and 223 (with zero emissions for bag 2) were used in the logarithmic regressions; could be included in footnote in the data analysis section. (Caretto)

c) It would be useful to show the confidence limits for the future prediction. (Caretto)

Appropriateness of the Data Sets Selected

a) Some explanation of the sample size used (77 as opposed to 78 or 80) would be helpful. (Gerritson)
b) The recruitment method (volunteers) may introduce sample bias. Not clear whether the sample is representative or whether the two subsamples (Ohio and Ann Arbor) are comparable, since the Ohio vehicles were recruited from I/M lanes and the Ann Arbor vehicles were not subject to I/M. Many of the Ohio vehicles were I/M failures, and therefore not a random sample. (Gerritson)
c) Reviewer not aware of any other data sets that could be used for this regression analysis. (Caretto)

Data Analyses Conducted

a) The use of regression analysis to assess the correlation between the FTP and HR505 results is appropriate. (Gerritson) The choice of logarithmic model is appropriate. (Caretto)
b) The high R^2 values in the regressions implies that the regression equations should be good predictors of hot running 505 emissions. (Caretto)
c) A simple regression (either linear or logarithmic) using only bag 3 as the regressor provides nearly the same value of R^2 as the regression equations presented in the report. (Caretto)
d) More discussion of the statistical analysis would be helpful. (Gerritson)
e) There may be some inconsistency in the rejection of model year data as insignificant, but the retention of insignificant regression parameters such as the bag 1 emissions for the logarithmic CO and NOx regressions. This inconsistency should not have any significant effect on the final results. (Caretto)
f) Two reviewers said it may not be appropriate to exclude vehicle #16. One reviewer said it may not be appropriate unless similar intermittent failures are known not to occur in the field. (Green) The other peer reviewer stated that it is not clear that excluding this vehicle had any impact on the results, and other vehicles with unrepaired problems were left in. (Gerritson) The third reviewer stated that the removal of vehicle #16 as an outlier appears to be a reasonable decision, judging by the erratic results for the different phases of the emission test. (Caretto)
g) Vehicle #219 also appears to be an outlier, at least for CO and HC emissions. (Caretto)
h) The data set has several vehicles with negative start emissions. Such data points should be kept in the analysis to avoid bias from removal of data points with negative errors without removal of corresponding data points with positive errors. Vehicles selected for the data set seem appropriate. (Caretto)
i) Since the average model year in the data set is 1991, the regression equations have better applicability for the later model years and technologies. (Caretto)

Appropriateness of the Conclusions

a) The conclusion is appropriate. (Gerritson, Caretto)

Recommendations

a) Consider testing another sample and comparing predicted values to actual values. Alternatively, recalculate the formula using only the Ohio data, then test it on the Ann Arbor vehicles. (Gerritson)
b) The three-bag form currently used to predict HR505 data should be compared to a model that excludes Bag 1, and a model that excludes both Bag 1 and Bag 2. (Green)
c) Statistical tests could be used to detect outliers, to determine if it is appropriate to drop vehicle #16 as well as vehicle #219 from this data set. (Caretto)

General Comments

a) The separation of exhaust emissions into running emissions and start emissions is a good choice. The definition of hot running 505 emissions and its use to compute start and running emissions appears to be the most logical choice. (Caretto)

M6.STE.003, "Determination of Start Emissions as a Function of Mileage and Soak Time for 1981-1993 Model Year Light-Duty Vehicles"

[Includes responses to issue #9]

Report Clarity

a) Specific task reasonably well described, but difficult to see how this report fits into the overall scope of the MOBILE6 emissions equations. (Caretto)
b) The paper is well organized and well written, but Appendix A is awkward and difficult to follow. (Gerritson)
c) Provide a reference for the high/normal emitter modeling concept. (Caretto)
d) The CARB report was introduced and used without a critical review. (Gerritson)
e) The discussion of the conversion from the CARB soak function to the EPA soak function is difficult to follow. Suggested text provided. (Caretto)
f) Section 3 should address real-world driving and the SFTP. (Green)
g) Section 4 should go to greater lengths to relate models to underlying physical phenomena. (Green)

Overall Methodology

a) The overall approach seems reasonable given the data sets that were available. (Caretto)
b) The definition of start emissions is appropriate. (Caretto)
c) The report does not discuss any details of the data, in particular, if any of the vehicles had negative start emissions and, if so, how they were handled. (Caretto)
d) No statistical results were presented, unlike the other reports. Expand Tables 4 and 5 to add regression statistics suggested by the reviewer. (Caretto)
e) In the conversion from the CARB soak function to the EPA soak function, two decisions were made which, while reasonable, were not explained. One was to have the EPA soak function equal to the California soak function for the entire soak time domain of the second empirical equation. The second was to use only a single value for the ratio of start emissions (10 min/720 min) for all control system types. (Caretto)

f) The procedure for determining the relative fractions of high and normal emitters is not discussed in this or any of the other three reports. The report states that the emission equations for normal and high emitters were estimated from FTP data, even though one of the terms in the equation is based on FTP data corrected by the Dayton IM240 results. (Caretto)
g) The assumption that the fraction of highs in start emissions is the same as the fraction of highs in running emissions is a good starting point for regime determination. (Caretto)

Appropriateness of the Data Sets Selected

a) The use of the FTP data set is appropriate for this analysis and the reviewer is not aware of any additional data sets that could be used. (Caretto)
b) It is difficult to comment on their appropriateness since no description of the datasets beyond their origin and responsible agency is provided. Also, no explanation of how the samples were chosen is provided. (Gerritson)
c) The use of the small API sample from Arizona is questioned based solely on the basis of odometer readings. (Gerritson)
d) The distribution of vehicles between Michigan and Arizona could be significant. (Gerritson)
e) The use of data from Ohio to correct for recruitment bias may not be appropriate. (Gerritson)
f) Reservation expressed about using two data points (start emissions for ten minute and twelve hour soaks) to adjust an entire continuum. (Gerritson)

Data Analyses Conducted

a) Correlation coefficients of less than 0.10 call into question the validity of the results, at least for high emitters. (Gerritson)

Appropriateness of the Conclusions

a) The implied conclusion appears to be appropriate, but it would be useful to have some measure of the statistical uncertainty in the final prediction equation. (Caretto)
b) The paper would benefit from a clear listing of the conclusions. (Gerritson)
c) The decision to divide the fleet into high and normal emitters is questioned. The high emitter category equates vehicles with emissions that may be several orders of magnitude apart, resulting in a scatter that has probably distorted the conclusions for this category. (Gerritson)

Recommendations

a) A procedure is outlined for checking the assumption that the fraction of highs in start emissions is the same as the fraction of highs in running emissions by comparing the high-running-emission vehicles with the high-start-emission vehicles. (Caretto)
b) Conduct further tests on high emitters, and develop a more robust statistic for this category. (Gerritson)
c) Conduct additional tests using soak times between ten minutes and twelve hours, to determine the validity of the adjustment factors. (Gerritson)

d) Redo the calculations using only the AAMA data, then apply the results to the samples from the EPA and API. (Gerritson)

M6.EXH.001, "Determination of Running Emissions as a Function of Mileage for 1981-1993 Model Year Light-Duty Cars and Trucks"

[Includes responses to issues #3 and 4]

Report Clarity

a) The paper is well organized, well written, thorough, and complete. Overall clarity is good. (Gerritson, Caretto)
b) On page 3, it would be helpful to explicitly relate the cold and hot start trip weightings to data collected to support the SFTP rulemaking. (Green)
c) Section 3.1 should clearly address the relationships between real-world driving, the FTP, and the SFTP. (Green)
d) Section 4 should include simple graphical comparisons of different models against a scatter plot of the data. (Green)
e) The equation near the top of page 5 does not readily follow from the preceding discussion. (Caretto)
f) On page 6, the description of case 2b does not fully specify how the final emission versus mileage relation is found. (Caretto)
g) The description of the derivation of the data for the high-emitter correction factor is not clear. The details of the "smoothing" analysis should be presented in the report. (Caretto)
h) The minor divisions for mileage on Figures 1 and 2, 4 and 2 kilomiles, could be more convenient if they were replaced by 5 and 2.5 kilomiles, respectively. (Caretto)
i) The note to Table 4 could be reworded to emphasize that the "adjustment additive" shown in the table has already been included in the slope. (Caretto)
j) Figures 1 and 2 would more clearly show a difference between the FTP and Dayton data sets if the 95% confidence intervals were shown for both the adjusted and the unadjusted curves. Alternatively, the data might be plotted to show the differences with a 95% confidence interval for the difference. (Caretto)

Overall Methodology

a) The overall method used to arrive at the running emission equations, including the data sets used, is appropriate. (Caretto)

Appropriateness of the Data Sets Selected

a) Overall, the datasets selected were appropriate, although the reviewer questions the use of the relatively small EPA and API datasets when the AAMA data alone would have been sufficient. (Gerritson)

Data Analyses Conducted

a) No statistical data are presented to justify the choice of the piecewise linear functions as compared with any other possible fits. Given the typical data scatter, however, there is no

apparent alternative to the piecewise linear approach that would give a significantly better representation of the data. (Caretto)

b) If the low mileage data are assumed to have zero slope, should they be used to develop a regression equation with a slope? (Caretto)
c) The reviewer questions the extent to which the piecewise linear approach represents a balance between simplicity and engineering judgment. The linear model presents the intuitive problem of ad infinitum deterioration. An example of a nonlinear form is given that appears able to do quite well. (Green)
d) In Section 5, it would seem more appropriate to adjust the data to account for demonstrated sample bias and then fit a physical model, than to fit a model to suspect data and derive post-fit "correction factors." (Green)
e) The correction factors based on the Ohio data may not be appropriate. While there may be sample bias, there are better approaches to deal with it. (Gerritson)

Appropriateness of the Conclusions

a) The conclusions for the normal emitters seem to be appropriate. Those for the high emitters are doubtful. The high emitter category equates vehicles with emissions that may be several orders of magnitude apart, resulting in a scatter that has probably distorted the conclusions for this category. (Gerritson)
b) It is appropriate to conclude that significant differences occur between the Dayton fleet and the FTP data fleet. The differences are consistent with the hypothesis that there is some recruitment bias in the FTP data fleet. (Caretto)
c) It is appropriate to adjust all the FTP data, regardless of the direction of the adjustment, not just the data for certain model years that are consistent with a recruitment bias, as EPA has done. (Caretto)

Recommendations

a) The problem of recruitment bias and the need to correct for high emitters should be dealt with in another study, designed to eliminate this problem. (Gerritson)
b) Test the hypothesis that the regression slope for low mileage is not significantly different from zero, to provide added support to at least one part of the piecewise linear regression equations. (Caretto)
c) It would be interesting to see whether eliminating the low mileage data from the unconstrained regression equation made any difference in the results. (Caretto)

M6.EXH.002, "Analysis of Emissions Deterioration Using Ohio and Wisconsin IM240 Data"

[Includes responses to issues #5, 6, 7, and 8]

Report Clarity

a) Although the report is well written and fairly easy to follow, it does not contain a clear exposition of the aims or objectives of the work. It would be helpful to have an abstract summarizing the purpose and results of the report. (Gerritson, Caretto)
b) The title of the report is misleading, since the actual deterioration analysis is presented in M6.EXH.001. (Caretto)
c) The report does not provide any quantitative justification for rejecting the corrected odometer data. (Caretto)
d) In Table 2, the variable named FSEC is not defined. (Caretto)
e) The report does not indicate if all the regressors shown in Tables 2 and 4 were used in the final regression equation. (Caretto)
f) In Table 4, one of the independent variables is erroneously listed as Lfxx instead of LxxIM. (Caretto)
g) Very little information was provided on the data set used to convert full IM240 emissions to estimates of running emissions. There was also no explanation for why the 59 vehicles from 1980-and-earlier model years were included in the regression intended for 1981-1993 model year vehicles. (Caretto)
h) The presentation of confidence intervals (such as is done for the unadjusted data in Figures 1 and 2) is a useful addition to the R^2 values. (Caretto)

Overall Methodology

a) A way to use the wealth of information generated by IM240 tests in estimating emissions deterioration appears to have become less viable as the study progressed. (Gerritson)
b) Other than the use of mileage accumulation data, the overall methodology is appropriate. (Caretto)
c) The conversion of fast pass IM240 measurements in Dayton to estimates of full IM240 appears to be a necessary step because there are no other data sets on late model year vehicles that provide measurements on vehicles that were not subject to I/M. (Caretto)

Appropriateness of the Data Sets Selected

a) How the vehicles required to complete the full test in Colorado were selected is not explained. This could be a source of bias. (Gerritson)
b) The selection of the state IM240 datasets and their inherent problems has introduced an unacceptable level of uncertainty into these analyses. There are too many confounding variables. (Gerritson)
c) The data sets used are generally appropriate. (Caretto)
d) There are no data to determine whether factors such as climate play a significant role in the representativeness of the Dayton data. (Caretto)

Data Analyses Conducted

a) The apparent problems with odometer readings underscore the need to seriously consider a time-based model. (Green)
b) The use of mean emission odometer readings ignores the distribution of mileage with vehicle age. (Caretto)
c) It seems quite a stretch to extrapolate, not just from Fast-Pass data to the full IM240, but ultimately to a full FTP or LA4. Since it appears to have a profound impact in some cases, the report should more clearly address and justify this use of the data. (Green)
d) The regression methods used here appear correct. (Caretto)

Appropriateness of the Conclusions

a) The report does not contain a presentation of the conclusions. (Gerritson)

Recommendations

a) The reviewer does not recommend that the currently available IM240 results be used at this time, given the uncertainties introduced through efforts to "adjust" these data. (Gerritson)
b) The FTP data could be analyzed to determine whether there is a difference in the emissions performance of vehicles as a function of location. (Caretto)
c) One possible analysis is to recompute the high emitter correction factor using the original (adjusted) odometer data. If this showed no significant difference in the results, the use of the mileage accumulation data would have greater justification. (Caretto)
d) One possible analysis is to use odometer and vehicle age data from other data sets. The alternate data could be represented as a set of empirical cumulative probability distribution functions for each vehicle age, which could be compared with the Dayton data to see how well those data matched the typical distribution. Comment intended to be exploratory in nature. (Caretto)
e) It would be useful to compare the RFF and EPA approaches on the same data set to determine which one provides the better correlation and report the results. (Caretto)
f) It would have been instructive to compare frequency distributions for the full IM240 results in the (measured) Wisconsin data with similar results in the (fitted) Dayton data. If these distributions were essentially the same, the Wisconsin data could be used directly. (Caretto)
g) Consider an analysis for MOBILE based only on age, using actual purchase dates to compute age where available, intended to be exploratory in nature. (Caretto)

General Comments Pertaining to the Four Reports

a) The overall approach taken in the four reports provides a reasonable method for the computation of start and running emissions of HC, CO and NOx for 1981-1993 light-duty vehicles. (Caretto)
b) An overview report describing the entire process for obtaining the start and running emissions would be helpful. (Caretto)
c) EPA staff should consider the preparation of an overall report to accompany the release of the

final version of MOBILE6 that would give an overview of the process used to develop the various components of the model. (Caretto)

d) The separation of FTP emissions into start emissions and running emissions is a good approach to the modeling of emissions from actual vehicle use. (Caretto)
e) The data sets used in these analyses appeared to be the most useful ones for the development of emissions for the 1981-1993 light-duty vehicles. No additional data sets are recommended for the emissions analysis. (Caretto)
f) The models are relatively simple, which is appropriate for fleet emissions that have a large scatter. (Caretto)
g) The presentation of the results from the regression analyses used in these reports is uneven. Having a complete set of regression statistics for all the equations used would be helpful. An informative statistic to report would be the confidence limits for the mean value of the dependent variable for a given set of independent variables. An estimate of the probable error in the final equations should also be provided. (Caretto)
h) There is no explicit conclusion section in three of the reports. (Caretto)
i) In general, it would be better to report the quantitative data used to justify conclusions made regarding alternative analyses. (Caretto)
j) Appropriate sections should specifically identify the physical processes relevant to emissions formation and control (e.g., misfire rate, catalysis, fuel characteristics, etc.), and explain how they lead to statistical model(s) for emissions deterioration. (Green)
k) EPA should offer clearer justification for relating deterioration to accumulated mileage, rather than either time or both time and accumulated mileage. (Green)

Comments on Issue # 1: Model Year Groupings and Segregation of Fuel System Types (applies to all papers)

a) The basis for the choice of technology groups is sound, although there are no analyses presented to justify the choices. (Caretto)