United States
Environmental Protection
Agency

Air and Radiation

EPA420-R-01-042
July 2001
M6.SPD.001

Development of Speed Correction Cycles

Printed on Recycled Paper

EPA420-R-01-042
July 2001

Development of Speed Correction Cycles

M6.SPD.001

Original Report by Sierra Research
April 30, 1997

Response to Comments, Corrections and Clarifications

Megan Beardsley
Assessment and Standards Division
Office of Transportation and Air Quality
U.S. Environmental Protection Agency

Development of Speed Correction Cycles
Response to Comments, Corrections and Clarifications

The Sierra Research report "Development of Speed Correction Cycles," SR97-04-01, April 30, 1997, was posted on the MOBILE6 web site as M6.SPD.001 in June 1997 for stakeholder comment. It was also sent for independent peer review.

While EPA used the original report to design driving cycles that were used to collect data for developing speed and roadway type corrections used in MOBILE6, review of the comments on the original report suggested the need for a number of corrections and clarifications to the report. Rather than revising the body of the completed contractor report, EPA has chosen to append a list of comments and corrections to the report.

We also have appended a complete list of comments received on the report and our response to those comments. These are divided into three sections: stakeholder comments, peer review comments from Dr. John Warner, and peer review comments from Professor H.Christopher Frey.

EPA Corrections and Clarifications

1. The title and body of the report refers to "speed correction cycles." While the results of testing on these cycles will be used to adjust emissions based on estimated average driving speed, these corrections are also meant to be used to correct for differences in driving on different roadway types ("facilities") under different conditions. Thus, it would be more precise to replace the term "speed correction cycles" with a term such as "facility- and speed-specific driving cycles."

2. On page 2 and again on pages 25-27, the report describes the development of an area-wide cycle that was constructed from several freeway segments and a non-freeway cycle. The report does not adequately explain the purpose of the area-wide cycles. These cycles were developed for eventual comparison with MOBILE6 area-wide results and for comparison with other area-wide speed cycles. The area-wide cycles have not been used for testing vehicle emissions. While the facility-and-speed specific cycles discussed in Chapter 3 were developed from the chase car data, the area-wide cycles were developed from the instrumented vehicles. Since the instrumented vehicle data did not distinguish arterial and local driving, a single "non-freeway" sub-cycle was developed.

3. On page 4, the report states that, in MOBILE5, speed correction factors are used to adjust the emissions measured using the Federal Test Procedure (FTP). More specifically, the

MOBILE5 speed correction factors are applied to the basic emission rate, a user-specified weighting of the three "bags" of the FTP.

4. The reference to "bags" on page 6 may be confusing. The city-specific area-wide emissions are simply a weighted average combination of various cycles as described on page 26 of the report.

5 On page 10, the report says that similar speed/acceleration frequency distributions for LOS A-C indicated that this range of LOS could be represented by one cycle. An additional reason for grouping these LOS together is that the chase car data indicates that these LOS are characterized by similar power frequency distributions. Since power-frequency distributions are highly correlated with emission differences, their consideration was an important factor in the decision to combine these LOS into a single cycle. Similarly, power-frequency distributions were considered in the decision to develop separate cycles for LOS D, E and F and to create a new "LOS G".

6. On page 12 there is a discussion of the "DiffSum" statistics used to compare driving cycles to the target driving population. "DiffSum" is calculated as the sum of the absolute values of the differences in the frequency distribution for each speed/acceleration bin.

7. On page 12, the report mentions that one criterion for evaluating cycles was the amount of operation in high specific power modes. The two modes (200-299 mph^2/sec and $\geq$300 mph^2/sec) were selected because EPA research conducted during the development of the Supplemental Federal Test Procedure (SFTP) indicates that in these ranges, some vehicles are designed for "commanded enrichment" which can significantly increase emissions. We felt it was important that these modes not be under- or over-represented in the new cycles.

8. In the report, "segment" has different meanings. In Chapter 3, "segment" refers to continuous driving data on the same roadway type and congestion level by a particular vehicle, as defined on page 11. In Chapter 4, "segment" refers to a driving cycle, in particular, a sub-cycle of the area-wide driving cycle. In Chapter 5, "segment" is used in both ways--the meaning should be inferred from the context.

9. The footnote on page 25 should read: "Large differences occur at <u>specific</u> congestion levels..."

10. The report does not make clear the units for the equations on page 26. Composite Emissions and Emissions are in g/mile. Weighting factors are a unit-less fraction. The speed/acceleration frequency distribution (SAFD) and travel fraction (TF) are both unit-less fractions. Note that, while the description of the equation for composite emissions says the weighting factor represents the fraction of travel "in miles," for convenience the

sample area-wide cycles described in this report actually use time-based weighting factors derived from instrumented and chase car data as described on page B-2. Similarly, time-based weighting factors were used for the Los Angeles area-wide weighted cycle described on page 33. If area-wide cycles were to be constructed for another city, distance-based weighting factors from local transportation model outputs would probably be used.

11. Page 26 mentions that travel mix was adjusted to account for "short trip bias." This adjustment is detailed in Appendix B.

12. The equation for the non-freeway speed/acceleration frequency distribution (SAFD) as printed near the bottom of page 26 is incorrect. The correct equation (and the equation actually used in this analysis for the calculation of the non-freeway SAFD) is:

$$SAFD_{NonFwy} = \frac{SAFD_{IV} - TFxSAFD_{FwyCC}}{1 - TF}$$

13. Page 28 incorrectly describes Figures 13 and 14, although the figures themselves are labeled correctly. Figure 13 is a three-dimensional SAFD (Watson Plot) of the Baltimore instrumented vehicle data. Figure 14 is a Watson Plot of the new area-wide cycle with weighting factors for Baltimore applied.

14. The information in Tables 7 and 8 on pages 32 and 33 comes from several different sources. The "population" values for facility-specific driving cycles in Table 7 are from 3-city (Baltimore, Los Angeles, Spokane) chase car data, while the population values for the LA-92 are from the Los Angeles chase car data only. In Table 8, the Non-Freeway Area-Wide segment is compared to the Baltimore weighted instrumented-vehicle data and the LA-92 is compared to the Los Angeles chase car data.

15. Page B-3 of Appendix B describes the calculation of a "bump-up" factor to ensure that travel fractions adequately represent short trips.. Because Los Angeles and Atlanta lacked instrumented and chase car data, respectively, the bump-up factors for these cities were calculated using data from Baltimore to replace the missing data. For example, it was assumed that the fraction of short trips in Los Angeles was equal to that observed in the Baltimore instrumented data. The Los Angeles bump-up factor was then calculated by dividing this assumed instrumented-vehicle short trip fraction by the actual (but under-represented) short trip fraction observed in the Los Angeles chase car data.

    Note, the Atlanta bump-up factor was calculated and listed in Appendix B for completeness but was never used.

Stakeholder Comments on M6.SPD.001 and EPA Response

Comments received from North Carolina Department of Transportation (NC DOT) (Comment #25), New York City Department of Environmental Protection (NYCDEP)(Comment #27), and American Automobile Manufacturers Association (AAMA) (Comment #37).

> *The following section reproduces stakeholder comments in plain text and intersperses EPA's response in indented italic. Commenters are identified by acronym at the end of each comment.*

Facility specific drive cycles are perhaps the most significant of the proposed improvements to the MOBILE Model. North Carolina approves of the move in this direction. The current model uses an average drive cycle to represent all possible driving conditions. This leads to counterintuitive results in some cases. However, we have some concerns based on the amount of aggregation and disaggregation in the supporting materials. Is the variability of stop/delay time implicit in the drive cycle that will be used to develop the basic emissions rates for each facility type? Our experience is that stop/delay time varies across facility types. We believe that any future version of the MOBILE model should account for this variation. An alternative method would be to allow the user to specify stop/delay time for each facility type. (NC DOT)

> *We agree that stop/delay time is important. The frequency of stops and delays were included in the speed/acceleration frequency distributions (SAFDs) that were used to develop the testing cycles, which were then used to develop MOBILE's facilty-specific speed correction factors. While users can not explicitly specify stop and delay time, they can specify differing fractions of VMT on different roadway types at different speeds–this effectively varies the stop/delay time included in the calculated emissions.*

We also note that arterials and collectors will share a driving trace. As noted above our experience indicates the existence of significant differences in stop/delay time and start mode between facility types. Collectors resemble locals streets more than arterial streets. (NC DOT)

> *Due to sample size and testing cost concerns, we were not able to develop cycles for all roadway classes at all speeds and LOS conditions. Instead it was necessary to group some roadway types together. Because chase car data (and hence, speed and LOS information) was not available for local roadways, grouping collectors with locals would mean that we could no longer distinguish collector driving by speed and LOS. For this reason we decided to group collectors and arterials. While there may be some differences between travel on these two roadway types, we believe that the ability to specify the average speed is sufficient for the most important distinctions between the two.*
>
> *Also, note that "start mode" is handled separately in the model and is not directly linked to roadway type. Varying start fractions by roadway type could be done but would require multiple runs of the model.*

The proposed freeway drive cycles also provided some surprises. The proposed drive cycles include: High Speed, LOS A-C, LOS D, LOS F, and LOS G. We recommend that the High speed drive cycle and the LOS A drive cycle be combined, and that the drive cycle for LOS B-C be kept together. Our understanding of the Highway Capacity Manual indicates that high speed driving occurs under LOS A. We also propose that LOS F and LOS G be combined. To the best of our knowledge, the Highway Capacity Manual does not recognize a LOS G. From the associated driving trace, this drive cycle represents breakdown conditions and might best be consolidated into the drive cycle for LOS F. (NC DOT)

It is true that our driving cycles do not match the Highway Capacity Manual. Our high speed cycle is a subset of driving under LOS A, B and C conditions, while our "LOS G" is a subset of driving under LOS F. These additional cycles were created because we needed to test vehicles at the low and high-speed extremes of highway driving. Tests at these points help determine the speed correction factors outside the range of more typical driving.

In particular, the development of the high speed and LOS G driving cycles and the grouping of the A, B and C driving cycles was partially based on the power frequency distributions for these cycles. LOS A freeway driving has a power frequency distribution similar to LOS B and C, while LOS G is signficantly different than LOS F and is useful for modeling the most congested conditions.

Have the new facility-specific cycles been reviewed by DOT/FHWA personnel? We are especially concerned about in-City roadways (arterials/collectors and local) where there are speed limits that may only allow 30-35mph. The maximum and average speeds for the bottom 3 cycles on Table 1, for congested in-City arterials and local roads, may be too high for many congested New York City streets during peak hours. New York City is likely to have a traffic control sign and signal density which is at the extreme end of the range in the nation. Frequency of starts and stops, and therefore of acceleration/deceleration, will not only affect average speeds but also the emissions associated with a given speed. We support any efforts by EPA to develop operating mode data that would allow us to project the impact on emissions of a high density of traffic signals on local streets as well as on arterials. (NYCDEP)

FHWA personnel have been involved in the discussions for how speed will be handled in MOBILE6 and have not expressed concerns about these cycles.

As will be described in the speed correction factor report M6.SPD.002, we have continued to test vehicles on the New York City Cycle and will use this data to project emissions at the low speeds that characterize New York City at peak hours.

Will speed corrections for arterials/collectors also utilize data from the NYCC and FTP cycles? Will the speeds on local streets be adjusted? If yes, what cycles will be used other than the

NYCC? How will idle CO emissions be calculated? Will they be calculated from the 2.5 mph emissions estimates? If yes, will the 2.5 mph emissions for local streets be the same as the 2.5 mph emissions for arterial/collectors? If they are not the same, how will they be calculated? How different can we expect the low speed correction factors to be in the new model compared to those used in MOBILE 5? Is there any reason why idle emissions data is not directly collected to use in the model instead of adjusting the 2.5 mph emissions? (NYCDEP)

These questions are to be addressed in M6.SPD.002, "Facility-Specific Speed Correction Factors."

This approach is somewhat different than the ARB's approach, in which a single, self-weighted inventory cycle was developed (the LA92, or Unified Cycle (UC)) and a significant number of cars and trucks were tested on this cycle. Also, ARB developed Unified Correction Cycles for developing speed correction factors for the UC. AAMA is unsure if EPA can devote enough resources to make their approach more accurate than ARB. Concern stems from EPA's desire to make one model fit all modeling purposes. The Unified Cycle approach is certainly more simple, and has the advantage that only a single, self-weighted cycle needs to be run for area-wide modeling. EPA's approach requires significantly more testing per vehicle, consequently, fewer vehicles can be tested. There is also an issue with respect to whether vehicles can be maintained at proper temperatures throughout the duration of the EPA cycle testing. AAMA recommends that EPA also have all of the vehicles tested on ARB's Unified Cycle as well as the other cycles, so the Unified Cycle can be compared to a weighted average of EPA's cycles. AAMA will reserve further comments on both ARB's approach and EPA's approach until it evaluates the data from EPA's test program, and particularly how EPA compares the data on the Unified Cycle to the data from the EPA's test cycles. (AAMA)

EPA has tested all vehicles tested using the EPA cycles on the LA92 (California Unified) cycle as well. A comparison of the LA92 results and the EPA cycle testing has been deferred until a final methodology for use of the EPA cycle testing is determined.

Significant steps were taken as part of the EPA cycle testing to minimize the effects of duration on emission results. All of the EPA cycles are much shorter in duration than the LA92 cycle. The longest cycle is 737 seconds (slightly over 12 minutes) long. No more than four cycles were run in series at one time and the order in which the cycles were done was randomized for each vehicle tested. Coolant temperature was monitered during testing.

EPA decided to pursue a facility based approach to evaluate the effects of driving on emissions when it became clear that an internally consistent method was necessary to allow the calculation of both area-wide inventories and the calculation of conformity scale emission analysis. EPA concluded that the incremental approach proposed by California was no more likely to succeed in satisfying both parts of the emissions estimate challenge (area-wide and conformity) than the approach taken by EPA. If properly conducted, both *methods should agree. Further, EPA believes that their approach will provide valuable insights into the need for and design of future testing programs and models. Future testing may require far fewer cycles than were tested for the current study, allowing for larger sample sizes.*

If EPA's approach of many factor-specific correction cycles remains unchanged for MOBILE6, it is essential that the model contain default (nationwide) statistics to develop average emission rates for a nationwide inventory. (AAMA)

National default estimates of VMT by speed and facility type will be available in MOBILE6. The development of these default values is described in M6.SPD.003, "Development of Methodology for Estimating VMT Weighting by Facility Type."

EPA must also allow users to output emissions based solely on current FTP certification test results, for ready comparison with the current and historical emission standards. (AAMA)

We considered an "FTP output" option when developing MOBILE6, but it was quite complicated to implement. We will keep this feature in mind for future versions of the model.

**EPA Response to Peer Review Comments from John Warner,
Center for Statistical Consultation and Research, University of Michigan, April 1, 1998**

The following document reproduces Dr. Warner's comments in plain text and intersperses EPA's response in indented italic. Note, this is a scanned version of Dr. Warner's comments. Not all mathematical symbols are correct.

This document contains a review of the report Carlson and Austin (1997) which was prepared by Sierra Research, Inc. for the United States Environmental Protection Agency. This review is divided into four sections. In Section 1, I give a brief synopsis of the report. In Section 2, I list eight specific recommendations for improving the statistical methodology of the report. In Section 3,1 give detailed explanations of each of the eight recommendations. In section 4, I list a few miscellaneous (that is, not specifically methodological) criticisms of the report. The review closes with the list of the technical reports on speed correction cycles that I have consulted in the process of conducting this review and a list of other references cited in the text.

In general, my conclusions are as follows: It appears that Carlson and Austin (1997) have done a good job in collecting and tabulating data from a variety of sources. The basic model (that is, the method for estimating area-wide emissions) is plausible and interesting. I am not enthusiastic, however, about the proposed method for extracting speed correction cycles from samples of speed traces. It seems likely to me that better methods can be found, particularly if a more specific criterion is given for judging the quality of a speed correction cycle. Finally, I feel very strongly that the statistical analysis of the speed trace data could be improved substantially. Indeed, Carlson and Austin (1997) present no inferential statistics of any kind, that is no p-values, no confidence interval, and no standard errors. I recommend that EPA obtain a new analysis of the speed trace data and I make a number of specific recommendations regarding the methods which should be employed in this analysis. In particular, the group or individual who carries out the new analysis should be familiar with random effects modeling and density estimation. The issues raised in Subsection 3.1 should be settled by EPA before this new analysis is attempted.

1 Synopsis of the report

Carlson and Austin (1997) present a new set of speed correction cycles (SCCs) for automobile emissions. SCCs are selected speed traces (or plots of speed against time). When automobiles are mounted on a dynamometer and run through a SCC, emissions are supposed to mimic the emissions of the population of comparable automobiles traveling on a road of a given type (facility type) under a given intensity of traffic congestion, or level of service (LOS). The model proposed by Carlson and Austin (1997) supposes that the aggregate driving conditions in all major US cities can be approximated (for the purpose of estimating auto emissions) by taking city specific weighted averages of SCCs from the proposed set of SCCs. If this method works as intended, one could accurately estimate emissions in all major US cities by 1) estimating the city specific proportion of travel on each facility at each LOS and 2) measuring the emissions on a representative sample of cars on all cycles in the set. Urban planners could then evaluate the effect of changing the mix of models in each city's passenger car fleet, or of changing the pattern of facility use, without undertaking any new emissions testing.

Carlson and Austin (1997) make use of three types of data to fit their model: speed traces sampled from chase cars (CCS), speed traces sampled from instrumented vehicles (IVs), and weighting factors obtained from network based transportation models. (A fourth type of data, laboratory emissions data from selected vehicles running along a specific SCC, would be used in future stages of model development). In the above, weighting factors are estimates of the proportion of all traffic in a given geographical location that travels on facilities of a given type at a given LOS. A population of speed traces is the set of all speed traces which occur for a given set of car/driver combinations occurring in a given geographical location under given restriction on facility type and LOS. (For example, the population of Ford cars driven by middle aged women on busy freeways in New York City). Chase cars can be used to sample from a population of speed traces as follows: First, a set of characteristic routes and starting times are chosen to represent the full range of driving patterns that obtain in a given city. Second, chase cars, outfitted with radar, are sent out to travel these preassigned routes at the appropriate times of day. While in transit, the chase cars are instructed to record speed traces from randomly selected vehicles. Instrumented vehicles (IVs) are randomly selected vehicles that are equipped with instrumentation for measuring and recording speed on a continuous basis. IVs produce samples of speed traces when they are driven by their owner's for a fixed period of time usually eight days (Defries and Kishan, 1992, p 3-21).

Carlson and Austin's (1997) report consists of 1) an executive summary and introduction, 2) a rough description of the methods by which SCCs are extracted from a sample of speed traces, 3) plots of the SCCs, 4) a description of the method for combining data of the four types describe above to estimate emissions in a given city, 5) a description of methods for evaluating the fit between a SCC and a corresponding sample of speed traces. A series of appendices is also provided. These contain tables of summary statistics, tabulations of empirical speed acceleration frequency densities (SAFDs) for various samples of speed traces and SCCs, and details concerning the calculation of area-wide emissions estimates. All this material is reviewed here. Carlson and Austin (1997) also presents a section on intersection analysis, which will not be reviewed because it is only tangentially related to the issues that I have been asked to consider.

2 Specific methodological recommendations

In this section I outline eight specific recommendations for improving the statistical methodology in Carlson and Austin (1997). Each specific recommendation will be discussed in a separate subsection below.

1. More care should be taken to define a specific set of summary statistics that can be used to determine when a sample of speed traces is a good representation for an underlying

population of speed traces. The same set of summary statistics should be used to determine when a SCC is a good representation of a sample of speed traces. Carlson and Austin (1997) use the SAFD for this purpose but Cohen, J. et al (1993) indicate that the SAFD is not adequate. This issue needs to be settled before the underlying statistical issues can be addressed in detail.

> *The U.S. EPA initially requested more detailed summary statistics from the contractor. However, given the time and resource constraints on this work assignment, we agreed to accept a more limited statistical description and comparison.*

2. Standard errors should be presented for the sample summary statistics discussed above. T-tests and ANOVAs (and their nonparametric analogues) should be used to determine if summary measures differ between populations. The unit of analysis for evaluating the summary statistics should be the estimated speed trace of an individual car/driver combination.

> *As mentioned above, we requested more detailed summary statistics from the contractor. However, given the time and resource constraints on this work assignment, we agreed to accept a more limited statistical description and comparison.*

3. A smoothing technique, such as kernel density estimation, should be used to estimate car/driver specific SAFD's.

> *We will consider using such techniques in the future when developing new cycles.*

4. Alternate methods should be considered for extracting a SCC from a sample of speed traces. I suggest (tentatively) a method for obtaining a SCC as a solution of a mathematical programming problem.

> *We will consider using such techniques in the future when developing new cycles.*

5. Claims made in the text of the report should be supported with summary tables including standard errors and p-values whenever these are available.

> *As mentioned above, we requested more detailed summary statistics from the contractor. However, given the time and resource constraints on this work assignment, we agreed to accept a more limited statistical description and comparison.*

6. Standard errors for the weighting factors should be obtained, if possible, from the network based transportation models.

We will consider this approach in the future when developing new cycles.

7. The equation defining $SAFD_{NonFwy}$ (near the bottom of page 26) is incorrect and needs to be modified.

   The equation on page 26 is listed incorrectly. We will make a note of this on the cover page accompanying the document. However, the actual computations used the correct equation as is clear from examining the non-Freeway SAFD which does sum to one.

8. Random effects modeling of car/driver specific SAFDs (or other summary measures) should be considered.

   We will consider this approach in the future when developing new cycles

3 Discussion of the methodological recommendations

This section provides a detailed discussion of the recommendations listed in the previous section.

3.1 Criteria for extracting speed correction cycles

An SCC is a short selected speed trace which is derived from, and intended to represent, an entire sample of (usually much longer) speed traces. More specifically, the average of any specific type of emission (per mile) for any car, traveling on a straight line with constant grade along a SCC, is supposed to be close to the average of the same emission (per mile) for the same car under identical conditions, traveling along the all speed traces in the sample. I will refer to this notion of agreement between SCC and sample as "total emissions agreement". This definition can be extended in an obvious manner to define a measure of agreement between two populations of speed traces, between a population and a sample of speed traces, etc. Carlson and Austin (1997) argue that their new SCCs will be representative of the samples from which they were derived by showing that the SCCs are close to the samples on a number of summary measures, (see pages 32-33). It appears that all of the stated summary measures are functions of the (sample or SCC) SAFD. This includes the PKE and the specific power distribution. As a result, agreement between the SAFD of the SCC and the SAFD the sample of speed traces implies agreement on all of the above summary measures. It is also apparent that the method for finding SCCs, proposed by Carlson and Austin (1997), explicitly attempts to minimize the difference between the SAFD of the SCC and the SAFD of the sample. Finally, Carlson and Austin (1997) display many pages of SAFD plots, apparently believing that these plots capture something fundamental about the relationship between the SCC and the sample. This leads me to define SAFD agreement between a SCC and a sample of speed traces as agreement between their underlying SAFDs. SAFD agreement between a population and a sample of speed traces can be defined in an analogous manner.

It is tempting to identify SAFD agreement with total emissions agreement. Such an identification, if it were permissible, would allow us to regard probablistic statements (i.e. confidence intervals) about estimators of the SAFD as if they were statements about corresponding SCCs. This is useful because, as I will show below, confidence intervals for estimators of the SAFD are easy to obtain. In addition, the identification of SAFD agreement with total emissions agreement has implications for the design of algorithms for extracting SCCs from samples of speed traces.

The crucial question now becomes: Is it permissible to identify total emissions agreement with SAFD agreement? I will refer to the hypothesis that the above two forms of agreement are (for all practical purposes) identical as the total emissions hypothesis. More specifically, I will define the total emissions hypothesis as the proposition that the instantaneous rate of emissions for any fixed car, moving in a straight line on a constant grade, is a function of speed and linear acceleration alone. In symbols:

$$e = f(s, a)$$

where e is the instantaneous rate of emissions, s is the speed, a is the linear acceleration, and f is an unknown function. There would seem to be good theoretical reasons for supposing that the total emissions hypothesis is, at least approximately, true. An argument for this would go as follows: Emissions result from the burning of fuel and as such they ought to be proportional to the rate at which the automobile dissipates energy. It would seem that energy is dissipated by 1) changing the car's kinetic energy, 2) overcoming drag, and 3) heat loss. It seems clear the first two of these energy sinks' are functions of speed and acceleration. It is less clear, however, that heat loss, particularly in the engine, could be fully accounted for by speed and acceleration alone. In fact, Cohen et al (1993), page 7, cites studies which indicate that speed and acceleration along cannot account for all emissions. I recommend that this matter be settled definitively (if it has not already been settled) either by equipping instrumented vehicles with emissions detection equipment (as in the studies cited by Cohen) or by means of computer simulation models using programs such as PC-VEHSIM, Carlson and Austin (1992). If the total emissions hypothesis is not even approximately true, then all of the evidence given in Carlson and Austin (1997) in support of the new SCCs, is suspect. If, on the other hand, the total emissions hypothesis is substantially true (as I suspect) then we should proceed to make full use of the SAFD in developing SCC extraction algorithms and in obtaining confidence regions.

3.2 Standard errors for the SAFD and summary statistics

Appendices A and C contain many tables of empirical SAFDs from various populations of speed traces. Summary statistics derived from these SAFDs are also presented. No standard errors or confidence intervals are presented for any of these parameters. This makes it impossible to judge the variability of these estimates. I recommend that standard errors for the above quantities be formed by treating each individual car/driver combination as an independent observation. Estimates of population wide summary measures can then be obtained by taking means of summary statistics calculated on each car/driver combination. Standard errors for these estimates can be calculated in the usual manner. For example, suppose one wanted an estimate the population SAFD at a point (s_0, a_0). I will call this quantity $S(s_0, a_0)$. To compute an estimate for S(s0, a0), begin by estimating the SAFD of each car/driver combination at (s_0, a_0). These estimates will be called $\hat{S}_i(s_0, a_0), i = 1,..., N$, where N is the number of car/driver combinations. The sample mean and standard error for $\{\hat{S}i(s_0, a_0), i = 1,..., N\}$ can then be used as a point estimate and standard error for $S(s_0, a_0)$. More complicated, and potentially more accurate, methods for estimating $S(s_0, a_0)$, could be developed by using random effects models (see below).

3.3 Kernel density estimators of car/driver specific SAFDs

I suggest that kernel density estimators be used to estimate the car/driver specific SAFDs (Silverman, 1986). These estimators will generally have better statistical properties then the histogram estimators used by Carlson and Austin (1997). Unlike histograms, kernel density estimators do not rely on arbitrary cutpoints for distinguishing between neighboring bins. Instead, kernel density estimators use an empirically determined bandwidth (or smoothness) parameter to produce estimates of a continuous density curve at any given point by taking weighted averages of observations found near that point. Kernel density estimators will outperform histograms most dramatically when one is attempting to estimate a density based on a relatively small number of observations. This is important because (as discussed in Subsection 3.2 I am recommending that, for IV data, a separate estimate of the SAFD be computed for each individual car/driver combination and, for CC data, a separate estimate of the SAFD be computed for each car/driver combination on each facility type at each LOS. Under these conditions some SAFDs may need to be estimated from relatively short driving segments. Kernel density estimators are available, and easy to compute, using the S-PLUS statistical software package (Statistical Sciences 1995). Some issues that arise in density estimation include the choice of an appropriate band width and the overcoming of edge effects, which result from the restriction that speeds may not be negative. A full resolution of these issues is beyond the scope of this review, and could not, in any case, be accomplished without access to the speed trace data. I recommend that the EPA contract with a competent statistical consulting organization to help with this project.

3.4 A mathematical programming algorithm for extracting a SCC from a sample of speed traces

The methods used by Carlson and Austin (1997) for extracting SCCs from a sample of speed traces has a number of apparent shortcomings. First, the method does not guarantee that an optimal fit (in any sense) to the sample SAFD has been found. Second, the method is subjective in the sense that it requires human intervention to edit prospective speed correction cycles. Third, the requirement, that the SCC be composed of segments of speed traces which actually occur in the sample, greatly complicates the process of finding optimal SCCs and is never justified rigorously. I suggest that the task of extracting a SCC from a sample of speed traces be formulated as a constrained optimization problem. To illustrate the process, suppose that we are given a smooth estimator of a SAFD, which we will call $\hat{S}(s, a)$. $\hat{S}(s, a)$, will be a density function defined for non-negative speed s and acceleration a. In practice, $\hat{S}(s, a)$ would be a (possibly weighted) average of car/driver specific estimators of SAFDs (of the form $\hat{S}_i(s_0, a_0)$, i=1 ,..., N, described in Subsection 3.2).

Our goal is to find a speed trace with a SAFD which is very close to $\hat{S}(s, a)$. I will suppose that this speed trace is to be 100 seconds long, and that acceleration is to be piecewise constant with jumps at each whole number time point. All speed traces of this kind can be represented by an ordered pair (V, A), where V is a scalar initial speed and A is a 100 dimensional vector with entries A_i, i = 1,..., 100. A_i will be the acceleration between time i - 1 and time i. I will denote the kernel density estimator of the SAFD of the speed trace (V, A) by $\tilde{S}(s, a)[V, A]$. Consider the measure of agreement between $\hat{S}(s, a)$ and $\tilde{S}(s, a)[V,A]$ given by

$$\int\int \hat{S}(s,a) - \tilde{S}(s,a)[V,A]^k ds da \qquad \textbf{(2)}$$

For fixed $k \geq 1$ we would like to find a value of (V, A) that minimizes (1). Side conditions like

$$|Ai - Ai - 1| < C, i = 2,...,100, \qquad \textbf{(3)}$$

for some $C > 0$, could also be imposed. The inequality in (2) insures that the acceleration does not change too rapidly. (Changes in acceleration that are too rapid may result in SCCs that cannot be driven). The integrand in (1) is differentiable for $k > 1$ and subdifferentiable for $k \geq 1$. If the integral in (1) is approximated by a double summation, the resulting optimization problem can be solved by standard methods. See Fetcher(1987) and NAG (1997) for a discussion of optimization algorithms which assume differentiability. Hiriart-Urruty and Lemarchechal (1993), and Schram and Zowe (1993), discuss bundle trust region algorithms, which apply to the subdifferentiable case. If we accept the total emissions hypothesis, as defined in Subsection 3.1, then an algorithm of the form described above may produce SCCs that are better matches to the target SAFD than those proposed by Carlson and Austin (1997) It is likely that the improvement would be most pronounced for freeway ramp SCCs, which appear to have poor fits to their target SCCs (see Table 7 on page 32 of Carlson and Austin (1997)). If, on the other hand, the total emissions hypothesis cannot be accepted, then the measure of agreement as given in (1) will need to be modified.

3.5 More and better summary tables

Arguments such as those presented in the first paragraph of page 25 need to be supported with additional summary tables. Standard errors and p-values should be included in these tables whenever appropriate. It is not sufficient to point the reader to the lengthy undigested computer output in the appendices with statements such as "By studying the detailed matrices the differences in freeway operations between the cities become apparent".

3.6 Standard errors for the weighting factors

Standard errors for the weighting factors, which appear in the definition of composite emissions at the top of page 26, should be obtained (if possible) from network based transportation models. Error propagation formulas can then be used, in conjunction with standard errors from laboratory emissions testing, to estimate the standard errors for the composite emissions. It may be possible to adjust the standard errors for the composite emissions to account for variability in the underlying SCC.

3.7 Defining $SAFD_{NonFwy}$

The equation that defines SAFDNonFwy near the bottom of page 26, is incorrect as it stands. It should read

$$SAFD_{NonFwy} = \frac{SAFD_{IV} - TFxSAFD_{FwyCC}}{1 - TF} \qquad \textbf{(4)}$$

This modification is required to insure that $SAFD_{NonFwy}$ integrates to 1. Even (3) needs to be used with care when estimates for $SAFD_{IV}$, TF, and $SAFD_{FwyCC}$ are used in place of their true population values (because the resulting estimate for $SAFD_{NonFwy}$ can assume negative values). I suggest that one uses an estimator which rounds negative values of SAFDNonFwy (in (3)) up to zero. $SAFD_{NonFwy}$ should then be renormalized so that it integrates to 1.

3.8 Random effects models

In some instances, the summary statistics from car/driver specific estimates of the SAFD may not be independent. For example, when one compares the SAFDs from chase car data across LOS groups, many car/driver combinations are likely to contribute to more than one LOS group. That is, the same car/driver will be followed across several LOS boundaries. In such a case, car/driver should be regarded as a random effect, leading one to consider mixed effects models. Mixed effects models can be fit using SAS PROC MIXED (SAS, 1997; Little, Milliken, Stroup, and Wolfinger, 1996). On a more abstract level, when one is estimating population-wide densities of the SAFD from car/drive specific SAFDs, it is intuitively clear that more weight should be given to estimates from car/driver combinations that have been observed for a longer period of time. To address this issue, one could use weighted least squares in the mixed modeling described above. Weighting each observation by the square root of the time observed might be a reasonable option. Estimating the standard error of the SAFD summary statistic would lead to better weights.

4 Other comments

1. A more detailed description of the methods used for calculating “bump up’ factors should be given. In particular, it needs to be explained how the equation in the middle of page B-3 could be used to compute “bump up” factors in the absence of data from both chase cars and instrumented vehicles. i.e. for Atlanta and Los Angeles.

   Because Los Angeles and Atlanta lacked instrumented and chase car data, respectively, the bump-up factors for these cities were calculated using data from Baltimore to replace the missing data. The Atlanta bump-up factor was never used.

2. The hypothesis that there are differences in speed traces for different classes of car, i.e. sports cars as opposed to economy cars, should be tested.

   The hypothesis that different kinds of cars are driven differently is intuitive, and in an ideal world we would test this assumption. However, given the time and resources available for testing, we think it is reasonable to treat all cars as an aggregate group, especially since we do not expect MOBILE users to have fleet and activity input data that specifies the “class” of car.

3. The hypothesis that $SAFD_{NonFwy}$ is city independent is stated as an assumption (see bottom of page 26). This should be tested.

   Graphical and statistical cross-city comparisons of driving on specific facility/LOS combinations are provided in Appendix A pages A-40 through A-41 and pages A-58 through A-67. While these comparisons are not a conclusive verification of the assumption that driving cycles are not city-dependent, they are consistent with this assumption. Given the time and resources available, we felt the assumption was a reasonable one.

Driving Cycle Literature Reviewed

Austin T et al (1997). "Methodology for Generating Driving Cycles for Inventory Development". Report No. SR95-09-02, Sierra Research Inc.

Austin T et al (1992). "Characterization of Driving Patterns and Emissions for Light-Duty Vehicles in California". Report No. SR92-11-02, Sierra Research Inc.

Carlson T and Austin T (1997). "Development of Speed Correction Cycles". Report No. SR97-04-01. Sierra Research.

Carlson T and Austin T (1994). "Evaluation of New Driving Cycles", Report No. SR94-0903 Sierra Research Inc.

Cohen, J. et al (1993). "Methods for Driving Cycle Development and Validation." Systems Applications International.

Defries T. and Kishan 5. (1992). "Light-Duty Vehicle Driving Behavior. Private Vehicle Instrumentation, Volume I", Technical Report. Radian Corporation.

Other Reference

Fletcher R (1987). Practical Optimization. Wiley New York.

Hiriart-Urruty, J-B. and Lemarchechal, C. (1993). Convex Analysis and Minimization Algorithms. Springer-Verlag, Berlin.

Little, R.C., Milliken, G.A, Stroup, W.W., and Wolfinger R. D. (1996). SAS Systems for Mixed Models. The SAS Institute, Cary, NC.

NAG Ltd. (1997). The NAG FORTRAN Library Manual, Mark 17. Programming. The Numerical Analysis Group, LTD., Oxford, UK.

SAS Institute (1997). The MIXED Procedure. SAS/STAT changes and enhancements through release 6.12 pp. 571-702, Cary, NC.

Schram, H., and Zowe, J. (1993). A version of the bundle idea for minimizing a nonsmooth function: conceptual idea, convergence analysis, numerical results. SIAM Journal on Optimization, 2, 121-152.

Silverman, B.W. (1986). Density Estimation for Statistics and Data Analysis. Chapman and Hall, New York.

Statistical Sciences (1995). S-PLUS Guide to Statistical and Mathematical Analysis, Version 3.3. StatSci, a division of MathSoft Inc., Seattle.

**EPA Response to Comments from Prof. H. Christopher Frey,
North Carolina State University, March 1998**

> *The following document reproduces Prof. Frey's comments in plain text and intersperses EPA's response in indented italic.*

The purpose of this review is to provide comments on the methodology and validity of the assumptions used in the report "Development of Speed Correction Cycles." The comments are organized based upon the major sections of the report.

Title:

The title does not accurately convey the content of the report. The report describes development of new facility-specific driving cycles, which is a significant departure from the speed correction cycle approach currently used. A better title might be "Development of Facility-Specific and Weighted Average Area-Wide Driving Cycles"

> *While the title has no impact on the conclusions of the study, clear writing is important to us. However, changing the report would require a new contractor work assignment. We will communicate this suggestion to the contractor as feedback and try to be clearer in other references to this work.*

Chapter 1 (Summary)

The use of facility-specific cycles that represent operation under different conditions is a significant advancement over the current speed correction cycle approach, and it holds promise for substantial improvement in highway vehicle emission factor estimation.

A key issue for future work is to assess whether the emissions for the new cycles are significantly different from previous cycles. Furthermore, it is also important to determine whether emissions on the new cycles are significantly different from each other. If two or more cycles produce similar emissions, then it may be possible to reduce the number of cycles used for additional vehicle testing. While not part of the scope of this report, recommendations should be made to EPA in this regard.

> *This is a useful suggestion for future work.*

The non-freeway segment of the area wide cycle (this is not a "speed correction cycle", so the name should be changed) shown in Table 2 has a higher maximum acceleration than the arterial and local roadway cycles shown in Table 1. Somewhere this should be explained. It is not clear why the arterial/LOS and local roadway cycles were not weighted in some manner as input to the development of an area-wide non-freeway cycle.

We will add a cover sheet explaining unclear references, including what is meant by "speed correction cycle." The question on area-wide non-freeway cycles is addressed below.

Chapter 2 (Introduction)

The first paragraph should be double-checked for accuracy. It is unclear from other documentation as to whether the MOBILE5 model is fundamentally based upon the entire FTP. cycle, or whether it is actually based on Bag 2 of the FTP cycle, which is the hot stabilized portion. The speed correction cycles mentioned in the first paragraph are all, as I understand it, in hot stabilized mode. Thus, there are other correction factors aside from that used for average speed, to adjust for differences in operating mode, ambient temperature, etc. Presumably the intent here is that all proposed cycles would also be in hot stabilized mode, although this could be an issue for the local cycle and some of the arterial cycles. This should receive some discussion.

We will add a cover sheet explaining unclear references, including the discussion of the Mobile5 model.

A key deficiency of the speed correction factor approach, which perhaps is implied but is not explicitly stated, is that it involves an "interpolation" from one average speed to another. However, since average speed is a meaningless representation for a driving cycle, as described in the report, the speed correction factor actually facilitates *extrapolations* to nonexistent situations for which no data were collected for any cycle.

While "average speed" may not be the best representation of a driving cycle, we do not believe it is meaningless. In particular, average speed and roadway type are likely to be the most detailed information available to modelers trying to characterize driving in their metropolitan area, thus it is important to describe driving cycles by their average speed as well as the roadway type on which such cycles may be driven. It is true that MOBILE often extrapolates from available data to situations for which no data is available. This is one of the basic purposes of the model.

In the second paragraph, it is probably not correct to say that it is assumed that the speed correction factors apply universally, at least not in all cases, in the sense that many of the speed correction factors are clearly intended to represent particular types of roadways under particular traffic conditions. For example, the high speed cycles clearly represent freeway driving under low congestion conditions. In practice, the Mobile5 model may be misused, but in that case the assumptions leading to misuse reside with the model users, not with the model itself.

While it is true that high speed correction factors only make sense for low-congestion freeway driving, the fact remains that MOBILE5 does not provide a mechanism for distinguishing typical driving on different types of roadways at the same average speed.

Page 6: in the description of Task 2 it is mentioned that city-specific bag weighting factors can be obtained from transportation planning model outputs. Without more detail given, the reviewer is highly skeptical of this claim. Transportation *planning* models typically do not produce the microscale data required to develop speed-acceleration frequency distributions. Transportation *simulation* models may be able to produce such output, but require considerable "calibration" and tweaking, and are probably not at a state of development or practice where they can be reliably used for this purpose.

The reference to "bags" on page 6 may be confusing and we will clarify this in our cover note for the report. The city-specific area-wide emissions are simply a weighted average combination of various cycles as described on page 26 of the report. Guidance for deriving area-wide emissions from transportation models is provided in the report M6.SPD.004, "Guidance for the Development of Facility Type VMT and Speed Distributions."

The discussion of Tasks on pages 5-7 is probably useful from a contractual viewpoint, but is distracting to the reader in terms of following the main ideas. A preview of the main activities of the project in the order that they are given in the following chapters would be more useful to the reader.

Clear writing is important to us. However, changing the structure of the report at this point would require a new contractor work assignment. We will communicate this suggestion to the contractor as feedback.

Chapter 3: (Facility -Specific Cycle Development)

A key assumption mentioned at the bottom of page 9 and top of page 10 is that driving data recorded for a given facility and LOS are not dependent upon the city in which the driving was performed. This assumption must be verified.

Given the time and resources available, we felt the assumption was a reasonable one. Graphical and statistical cross-city comparisons of driving on specific facility/LOS combinations are provided in Appendix A pages A-40 through A-41 and pages A-58 through A-67. While these comparisons are not a conclusive verification of the assumption that driving cycles are not city-dependent, they are consistent with this assumption.

When deciding upon grouping of LOS categories, it is mentioned that LOS A, B, and C for freeways had similar speed/acceleration profiles. However, the more relevant criteria would be whether emissions differ significantly for these LOS categories. More justification needs to be given as to why LOS A, B, and C were sufficiently similar that it was judged that *emissions* would be similar.

Similarly, it was decided not only to keep LOS D, E, and F as separate cycles, but to add a fourth LOS "G" cycle. Even though these cycles have different speed and acceleration profiles, they may or may not have significantly different emissions. As an aside, we have found that in some cases the emissions for the low speed cycles used for the speed correction factor in Mobile5 are not significantly different (Kini and Frey, 1997).

An added cover sheet will explain unclear references, including the selection of LOS categories for cycle development.

The discussion of the High-Speed Freeway Cycle appears to be similar to that for LOS-Based Freeway cycles, with the exception of a minimum speed cut-off. Some discussion of speed ranges should be given for the LOS-based Freeway cycles for LOS A, B, and C to clarify the differences of this cycle compared to the high speed freeway cycle.

Table 3 on page 13 provide a summary of the speed and acceleration characteristics of the two cycles. Furthermore, the driving traces in Figures 1 and 2 provide a graphical depiction of the differences in the two cycles. In particular, the high-speed cycle lacks the low-speed driving of the LOS A-C Cycle and includes less acceleration.

The discussion of creation of three arterial/collector cycles on page 10 is not well-motivated. As discussed in the introduction, average speed is not a good way to represent a driving cycle. Thus, it is not a sufficient rationale for grouping data. Most likely, the discussion in the text is not fully representative of the thinking that went into development of the groupings, and additional discussion would solve this problem.

As noted above, facility-specific average speed is an important driving cycle characteristic because it is one of the few cycle characteristics that can be determined for local areas. Average power for each LOS was also considered, but the LOS groupings were based primarily on average speed and sample size (e.g., there weren't many LOS F observations for arterials and collectors, so sample size constraints led E and F observations to be grouped together).

The approach to using "segments" seems reasonable, but does raise issues regarding possible autocorrelation of emissions with the speed/acceleration history of the vehicle, which could lead to some unintentional discontinuities when segments are spliced together in terms of trends with which speed and acceleration may be changing at the end of one segment and the beginning of the next. However, the segment approach seems to be more realistic than the microtrip approach previously used.

The procedure used for splicing segments together, based upon matching of endpoint speed and accelerations of one segment to those of the adjacent segment, appears to be reasonable. However, the tolerances used (plus or minus 0.5 mph for speed and plus or minus 0.5 mph/s for acceleration) appear to be arbitrary unless they are more fully discussed. If there is some basis for these tolerances, they should be explained. For example, it would appear that 0.5 mph change in speed from one second to the next is reasonable, since this implies an acceleration of only 0.5 mph/s. A change in acceleration from, say, 0 mph/s to 0.5 mph/s implies a change in speed of only about 0.5 mph from one second to the next. However, if one segment has rapidly changing accelerations, and is spliced to a segment with constant acceleration, this might lead to a discontinuity that could influence emissions. Thus, another criteria to consider would be the rate of change in the acceleration at the end points of the segment.

We agree that there are situations where discontinuities between segments could theoretically create cycles with emissions different than the emissions from the underlying segments. While it is not clear that such discontinuities have a significant impact on the cycles developed here, matching the rate of change in acceleration at the end points of the segment would be helpful in avoiding these situations in the future, and we will consider including this criterion when developing future cycles.

In the meantime, the selection of 0.5 mph and 0.5 mph/s as tolerances for matching was based on engineering judgement since there was limited second-by-second data to determine exactly what tolerances would be best.

The description of "DiffSum" needs to be clearer. Presumably, the absolute values of the differences are used. Why not use a square error? It would help to show an equation for DiffSum so that there is no ambiguity regarding how it is defined.

DiffSum will be described in the cover sheet.

One potential problem with the DiffSum approach is that it is equally sensitive to many small deviations and to a few very large deviations. For example, if there were 10 bins of data and each was in error by one unit, then DiffSum would be 10. However, if eight of the bins were exactly correct, but two were in error by 5 units each, DiffSum would also be 10. It is likely that many would prefer the former case to the latter case, since the former case would seem to have a small random error while the latter case has a large error for some cases (perhaps indicative of a systematic error). If a squared error approach were used, the latter case would receive a much higher squared error (50) than the former case (10) and hence would be penalized more heavily.

The DiffSum approach may also depend upon the size of the bins used for the SAFDs. The sensitivity of the DiffSum statistic to bin size should be discussed.

These are useful criticisms of the absolute value approach to comparing cycles. In future work we will seriously consider using a squared value approach instead. However, we do not believe that the difference between the two approaches is likely to have significantly skewed the selection of segments for the cycles developed here.

An explanation of the emissions significance of the high specific power modes that are mentioned on the bottom of page 12 would help. Specifically, on what basis are emissions expected to be significantly different for the range of 200-299 mph^2/s versus ≥ 300 mph^2/s?

These two modes (200-299 mph^2/sec and ≥ 300 mph^2/sec) were selected because EPA research conducted during the development of the Supplemental Federal Test Procedure (SFTP) indicates that in these ranges, some vehicles are designed for "commanded enrichment" which can significantly increase emissions. We felt it was important that these modes not be under- or over-represented in the new cycles. We will add a cover sheet explaining unclear passages, including a more detailed description of the choice of high specific power modes.

It should be noted that the speed/acceleration distribution may not be the only main consideration that affects emissions. It is possible that autocorrelations (or the time series of speed/acceleration) are important as well. This possibility should be acknowledged and recommended for investigation when emissions data are collected for these cycles. To the extent that the times series or time history of speed and acceleration influences emissions, the procedure used here may require modification in the future.

The second-by-second emissions data collected on the cycles described here does allow the analysis of autocorrelation effects. While modeling such effects is beyond the scope of MOBILE6, current work elsewhere to develop"modal models" is beginning to investigate the time history of emissions, particularly for catalyst effects.

The description of the basic characteristics of each cycle can encourage some of the pitfalls of the previous speed correction cycle approach. It would be useful to move away from average speed as a description of a driving cycle. In Table 3 information for average speed, maximum speed, and maximum acceleration are presented. However, it would be more useful to know something about the joint distribution of speed and acceleration. While it may be easier to show this type of information graphically, using an approach such as in Figure 13, it would also be possible to create some large bins (each with a speed range and an acceleration range) and to indicate what portion of the total cycle time is contained in each bin. Even if only four bins were given (e.g., low speed and high speed, each with low acceleration and high acceleration) this might give a more meaningful indication of each cycle. Such information could be tabulated. Alternatively, information regarding the standard deviation of speed and acceleration, and regarding correlation among speed and acceleration (or speed and positive accelerations) would be useful.

If it is desired to use average speed to describe a cycle, then also include the standard deviation of speed so that one can gain some idea of the variability in speed associated with the cycle.

We initially requested more detailed summary statistics from the contractor. However, given the time and resource constraints on this work assignment, we agreed to accept a more limited statistical description and comparison.

Chapter 4: (Area-Wide Cycle Development)

In Chapter 3 cycles are presented for Arterial (LOS A-B, C-D, and E-F), and local roadways. These are not mentioned in Chapter 4. Why? Why is it not possible to develop an area-wide cycle by weighting of the freeway cycles with these additional cycles for arterials and local roadways? Or, stated more specifically, why was a single non-freeway cycle developed instead of weighting the arterial and local cycles? If a single non-freeway cycle is to be used, why were arterial/local cycles developed?

> *The report does not adequately explain the purpose of the area-wide cycles. These cycles were developed for eventual comparison with MOBILE6 area-wide results and for comparison with other area-wide speed cycles. While the facility-and-speed specific cycles discussed in Chapter 3 were developed from the chase car data, the area-wide cycles were developed from the instrumented vehicles. Since the instrumented vehicle data did not distinguish arterial and local driving, a single "non-freeway" sub-cycle was developed. This explanation will be added to the cover sheet*

What is the precision and accuracy of the measurements for speed made using the chase car? Are measurement errors negligible? Was acceleration calculated using a finite difference based upon second-by-second speed data? What is the error in the acceleration data/estimates? Some discussion of this is needed.

> *Sierra Research Report SR92-02-01, "Design and Operation of an Instrumented 'Chase Car' for Characterizing the Driving Patterns of Light-Duty Vehicles in Customer Service," prepared for EPA, February 29, 1992, describes the chase car methodology; however, the report does not quantify measurement errors. Acceleration is calculated each second from successive speed measurements.*
>
> *Speed estimates of instrumented vehicles are described in EPA report 420-R-93-007, May 1993, "Federal Test Procedure Review Project: Preliminary Technical Report." Many instrumented vehicles have a speed resolution of only 1 mph.*

In Chapter 3 a key assumption mentioned was that it has been assumed that driving on a particular facility and a particular LOS is similar in any of the cities. Has that assumption been verified by comparing chase car data from different cities for the same facility type and same LOS? This pertains to Table 4, in which it is assumed that it is possible to compare facility/LOS from one city to another. There is a brief mention of this on page 25, but more discussion is needed.

> *A similar comment was made on Chapter 3. See our response above.*

Table 4 is somewhat confusing because of all of the percentages shown. It would be helpful to show check sums of 100% at the bottom of each set of categories merely for clarity, or to separate this into three different tables. Presumably, the middle set of rows (labeled by LOS only) is for driving on the given LOS on all facility types. However, since data for arterials were developed based upon LOS designations, it would be useful to also show the breakdown of Arterial Total by LOS categories.

We agree this table could be clearer. However, given the resource constraints on this project, at this time we have decided not to ask the contractor to revise the table or to generate the travel distributions for arterials by LOS.

The discussion in the paragraph on the middle of page 24 would be strengthened by mentioning not only average speed, but also standard deviation of speed for Spokane and Los Angeles.

As mentioned above, we initially requested more detailed summary statistics from the contractor. However, given the time and resource constraints on this work assignment, we agreed to accept a more limited statistical description and comparison.

The presentation of Table 5 and the association discussion on pages 24 and 25 is difficult to follow. It might be more meaningful to show 3D graphs of the joint distributions for speed/acceleration for at least a few cases, so that the reader can gain some insight into what the "difference" represents. For example, a 3D graph of the joint distributions for speed/acceleration for each of LA and Baltimore, accompanied by a 3D graph of the distribution of "differences" between the two, would clearly indicate what the values in Table 5 represent. All three graphs could be placed on one page. Presumably the data used to develop Table 5 are based upon all LOS categories for freeways and non-freeways, respectively.

We would like to make the report easier to follow. However, the changes suggested here would require a new contractor work assignment. We will communicate this suggestion to the contractor as feedback.

The footnote on page 25 seems to be missing some words. It states "Large differences occur at the congestion levels..." Which congestion levels? Is this meant to refer to a particular LOS?

The footnote should read: "Large differences occur at <u>specific</u> congestion levels..." This will be noted in the cover sheet.

The word "segment" is used on the last paragraph on p. 25, but appears to have a different meaning than on pages 11-12. The terminology should be more clearly defined. Would it be correct to say that the area-wide cycle is comprised of facility/LOS-specific subcycles?

We will add a cover sheet explaining unclear passages, including the use of the term "segment."

The equation on the top of page 26 appears to be missing the LOS "G" cycle for freeways.

No, the area-wide cycle does not have an explicit LOS G sub-cycle. LOS G is included in the area-wide cycle as a subset of LOS F.

The weighting factors are said to be the fraction of travel in miles. This presumes that the emissions are in units of g/mi? The units are not defined. Please make sure units are correctly specified for the weighting factors, emissions, and composite emissions.

We will add a cover sheet explaining unclear passages, including units for the equation on page 26.

On page 26, it is stated that "short trip" bias was accounted for by adjusting the travel mix. How was this adjustment made? The answer is in Appendix B but some description of the approach would help. It is especially useful to mention that the adjustment procedure is based upon measured values from both the instrumented vehicles and the chase car data sets.

We will add a cover sheet explaining unclear passages, including a reference to Appendix B.

Why would the non-freeway target (speed/accel?) driving distribution be city-independent? Was the back-calculated non-freeway driving distribution compared to the non-freeway driving distribution obtained from the chase car in different cities to check for expected differences or expected similarities?

Because the chase car data underrepresents short trips, it was not possible to make a direct comparison of the non-freeway driving SAFD with the chase car data from the different cities. And, because the area-wide cycle is meant only for comparison and not for city-specific emission calculations, the assumption that the non-freeway driving distribution is city independent is appropriate for MOBILE6 purposes.

On top of page 27, more discussion is needed regarding the approach to development of the non-freeway area-wide cycle. It is stated that this cycle is based upon instrumented vehicle (IV) data. However, earlier (on p. 9) it was stated that facility type could not reliably be inferred from the IV data. Thus, what quality assurance was performed to assure that the data used from the IV data set truly represents non-freeway facilities?

The DiffSum procedure was used to select instrumented vehicle segments that represent non-freeway cycles. These are the segments that best match the characteristics of the non-freeway driving. Thus, while there is a small chance that some of these segments were actually driven on freeways, the segments match the speed and acceleration profiles of non-freeway driving and can be expected to produce a good representation of non-freeway emissions.

When comparing figures (i.e. Figures 13 and 14) it would help to have both on the same page or on the same spread (facing pages).

We would like to make the report easier to read. However, the change suggested here would require a new contractor work assignment. We will communicate this suggestion to the contractor as feedback.

Presumably, the cycle presented in Figure 13 is a weighted area-wide cycle for Baltimore that includes the four freeway LOS cycles, the freeway ramp cycle, and the non-freeway cycle. If so, it would help the reader to have more information regarding the specific example for Baltimore. Specifically, it would help to show the weighting factors used for each of these six constituent cycles. Note that in Appendix B time-based weights are given, whereas on page 26 distance-based weights are given in the equation for emissions. The discrepancy between these two needs to be corrected or explained.

While the equation for composite emissions on page 26 says the weighting factor represents the fraction of travel "in miles," for convenience, the sample area-wide cycles described in this report actually use time-based weighting factors derived from instrumented and chase car data as described on page B-2. Similarly, time-based weighting factors were used for the Los Angeles area-wide weighted cycle described on page 33. If area-wide cycles were to be constructed for another city, distance-based weighting factors from local transportation model outputs would probably be used.

The text on page 28 is incorrect. Figure 13 is the Baltimore instrumented vehicle data. Figure 14 is a weighted area-wide cycle. This correction will be noted in the cover sheet.

Chapter 5: (Evaluation of Driving Cycles)

There is a general problem in this report with the flow of ideas from Chapter 3 to Chapter 4 to Chapter 5. Chapter 4 almost seems to be a separate report on a different topic, since it deals with a non-freeway cycle and seems to ignore the arterial/local cycles. Chapters 3 and 5 appear to deal with the same set of cycles, but Chapter 5 includes the non-freeway cycle at the end. More transition is needed from one chapter to the next, since these changes in focus are confusing to the reader.

We would like to make the report easier to read. However, the change suggested here would require a new contractor work assignment. We will communicate this suggestion to the contractor as feedback.

The comparison of facility-specific cycles to the "population" is interesting and informative. However, the first comparison made is for average speed. As previously noted in the report, this is not the most important figure of merit when comparing cycles. Thus, it should be emphasized that not only are there minor differences in average speeds when comparing cycles with the population data, but that average speed is not by itself a meaningful descriptor of a driving cycle. It would be useful to include standard deviation of average speed and standard deviation of acceleration when comparing the cycles and populations. It would also be useful to consider factors such as correlations or covariance between speed and acceleration.

The discussion of maximum speed and acceleration is too brief. While it is not surprising that any particular sample from the population fails to capture the maximum value from the population, it would be useful to have some idea of how close the sample is to the population. For example, one could calculate the percentile of the population distribution for speed corresponding to the maximum value for the cycle maximum speed (e.g., is the cycle maximum speed at the 98th percentile of the population distribution for speed, or is it at the 63rd percentile?). If the maximum speed in the cycle were at a low percentile of the speed distribution from the population, that would suggest a possible inadequacy in the cycle, whereas if it were at a high percentile then there would be increased confidence that the cycle is representative of the population. Without this type of information, the comparison of maximum speed and maximum acceleration between the cycles and populations is not very helpful.

> *As mentioned above, we initially requested more detailed summary statistics from the contractor. However, given the time and resource constraints on this work, we agreed to accept a more limited statistical description and comparison.*

It would also be useful to explain a few cases. For example, for Arterial LOS A-B a maximum acceleration of 14.9 mph/sec is indicated in the population, but the maximum used in the cycle is only 5 mph/sec. Was the population maximum due to only one vehicle, with the next highest acceleration closer to 5 mph/sec? Or is there a large fleet of high performance vehicles that are not represented by the driving cycle.

> *The SAFD tables in Appendix C provide the speed and acceleration distributions for the target populations. Clearly, the driving cycles do not reproduce the most extreme accelerations and speeds. Significant additional analysis would be required to evaluate how many vehicles are responsible for the extremes in the population distributions. However, for emission estimates, the number of vehicles driven at these acceleration and speed extremes is less important than the fraction of miles driven.*

What was the LA92 cycle compared with? It is stated that the LA92 cycle was compared with its target driving population-- what population was actually used in the comparison? This is important since the total SAFD difference and high power difference for the LA92 are used as figure of merits for evaluation of the other cycles. It is important to clearly convey that the LA92 cycle is not being used as a strawman, but that the comparison of LA92 to the “target” population is fair and realistic.

> *The “population” values for facility-specific driving cycles in Table 7 are from 3-city (Baltimore, Los Angeles, Spokane) chase car data, while the population values for the LA-92*

are from the Los Angeles chase car data only. In Table 8, the Non-Freeway Area-Wide segment is compared to the Baltimore weighted instrumented-vehicle data and the LA-92 is compared to the Los Angeles chase car data. Since the LA92 cycle was designed to represent driving in Los Angeles, these comparisons should be favorable to the LA92. This information will be repeated in the cover sheet.

At the bottom of page 33 it is stated that "segment" (again, this word is used in different ways in different places in the report) weighting factors for Los Angeles were used to develop an area-wide cycle for Los Angeles. How were these factors developed? What were they? They should be summarized in a table.

The use of the term "segment" is described in the cover sheet. The weighting factors for Baltimore, Spokane and Los Angeles are explained in Appendix B and listed on page B-2. As described earlier for Baltimore, the time-based weighting factors were used.

The comparison of the Area-Wide and the LA92 cycles with the Chase Car data might be presented in a different order. It seems useful to compare LA92 with the chase car data, and the area-wide cycle with the chase car data. These two comparisons illustrated that the area wide cycle compares more favorably with the chase car data than does the LA92 cycle. The third comparison, of areawide vs. LA92, seems unnecessary and potentially confusing, but if it is desired to include it, make it the last one.

We would like to make the report easier to read. However, the change suggested here would require a new contractor work assignment. We will communicate this suggestion to the contractor as feedback.

For all of the comparisons, a key point to mention is that whether the error measures (total SAFD % difference and high-power difference) are useful or appropriate will not be fully known until emissions data are collected for these cycles. If there is any previous work in which these two error measures were used and found to be good predictors of agreement or lack thereof with respect to *emissions*, it would be useful and important to note that and discuss. In the absence of such empirical verification of the utility of these error measures, at best these error measures can be said to be based upon hypotheses regarding factors that are most influential with respect to emissions, and that new error measures may be proposed after emissions data are collected and evaluated. If new error measures are later proposed (based upon an empirical foundation), then new cycles may have to be developed based upon those measures.

The contractor who prepared this report also did analysis for EPA's Supplemental FTP study. The study found these error measures to be good predictors of emission agreement. Therefore, they continued using them in this cycle development work.

Chapter 6 (Intersection Influence Analysis)

Signalized intersections are of great importance. However, some discussion of non-signalized intersections (e.g., four way stops with stop signs) would be useful, at least for context. If any data are readily available on the number of signalized intersections in the U.S. versus non-signalized intersections, this would provide some context.

As stated on page 35, intersections with stop signs are considered signalized intersections. We do not know of readily available data on the fraction of intersections that are non-signalized.

Some more transition is needed, since the third paragraph refers to second-by-second data from speed correction cycles. This catches the reader off-guard, since the report states early on many of the problems with such cycles and then proposes new cycles.

We would like to make the report easier to read. However, the change suggested here would require a new contractor work assignment. We will communicate this suggestion to the contractor as feedback.

In the 2nd full paragraph on page 36, in the last sentence, there is mention of "new speed correction cycles." This is a misnomer, since the new cycles are not intended as speed correction cycles. They are intended for use in developing area-wide cycles based upon weighting of multiple cycles. The text in this paragraph should be revised.

We will add a cover sheet explaining unclear references, including what is meant by "speed correction cycle." However, it should be noted that the cycles discussed here are not meant to only be used in developing area-wide cycles. They are also used to simulate driving at specific average speeds on specific roadway types.

Since only a sample of data were used to create the information in Tables 10 and 11, it is important to make a notation regarding the total amount of driving time and driving distance upon which the fractions shown in the tables are based, especially for Table 11. In addition, for each a case a note should be made regarding the percentage of the actual population that was reviewed (there is one example of this for local roadways, but it is not stated whether the percentage is on a time or distance basis).

The distance and percentage numbers requested are not readily available. For this reason and because the intersection analysis was not used in MOBILE6, we have decided not to follow up on this comment at this time. However, if the results of this analysis are used at a later date, it would be worthwhile to ask the contractor to generate the requested data.

Since the comparisons of the cycles to the target population are of interest, it would be helpful to present the data in a form that facilitates comparisons. A landscape table with two columns for each cycle (one for the cycle and one for the population) would help.

> *We would like to make the report easier to read. However, the change suggested here would require a new contractor work assignment. We will communicate this suggestion to the contractor as feedback.*

Other Comments

A summary should be given of the recommended uses for the new facility/LOS-specific cycles, and for the areawide cycle approach, and for emissions data collection.

> *We would like to make the report easier to read. However, the change suggested here would require a new contractor work assignment. We will communicate this suggestion to the contractor as feedback.*

The report should conclude with a brief section highlighting the key assumptions underlying the development of the new cycles, and making recommendations for future work pending the outcome of emissions testing with the new cycles. Two key issues are:

- Whether there are significant differences in emissions among the new cycles. If two or more cycles have similar emissions, then it may be possible to reduce the number of cycles used for large scale testing efforts.

> *Now that we have emission results from these cycles, we agree that a statistical comparison of emissions on the new cycles would be a useful tool in selecting which cycles should be used in future vehicle testing.*

- The adequacy of the objective function/error measures used in constructing the new cycles. If the SAFD % Difference and high-power difference measures, and any others used, are found to not adequately capture sensitivities between speed/acceleration and emissions, then new cycles may have to be developed based upon new evaluation approaches.

> *We agree that tests of these correlations would be useful. Currently we are not planning to develop any new cycles, but we will keep this suggestion in mind for the future.*

Another factor to consider in emissions testing would be road grade, which is not discussed in the report. However, engine power demand and emissions could differ for the same speed trace depending upon the road grade. If road grade data are available from the three city study, this information would be useful to include as part of recommendations for new driving cycles.

> *The effect of road grade on emissions is beyond the scope of this study and the MOBILE6 model. However, we are considering ways to include road grade in future research.*