Validity of Chase Car Data Used in Developing Emissions Cycles

JENNIFER E. MOREY

THIRAYOOT LIMANOND

Institute of Transportation Studies

University of California, Davis

DEBBIE A. NIEMEIER

Department of Civil and Environmental Engineering

University of California, Davis

ABSTRACT

In an effort to ensure vehicle compliance with U.S. air quality policies, driving cycles, profiles of average driving behavior, have been constructed to characterize the driving behavior of the overall fleet. The cycles are built from chase car data, speed-time profiles of in-use vehicles recorded using a chase car method. This study evaluates the acceptability of using chase car data as the foundation for driving cycle development and recommends changes in the current data collection protocol. Two data issues are closely examined: 1) the effectiveness of the current target vehicle selection procedure and 2) the validity of blending data collected from target vehicles with data collected from the chase car, a method used when target vehicles are unavailable. Although in the aggregate there do not appear to be significant discrepancies between these chase car and target vehicle data, when examined at disaggregate levels, significant differences appear that could affect the representativeness of existing driving cycles. Recommendations include increasing the proportion of target to chase car data in future databases by improving the existing protocol and considering the use of different recording technology.

Debbie A. Niemeier, Department of Civil and Environmental Engineering, University of California, Davis, One Shields Avenue, Davis, CA 95616. E-mail: dniemeier@ucdavis.edu

INTRODUCTION

Driving cycles, profiles of average driving behavior, are used to certify new vehicles, to verify vehicle compliance with inspection/maintenance (I/M) programs, and to create emissions factors for performing transportation conformity determinations. Although there may be no single representative driving cycle, characterizing average driving behavior is a very important element in describing overall fleet emissions. Numerous data have been collected to create these driving cycles. To date, two data collection methods have most often been employed: 1) the use of a chase car to mimic driving behavior while recording speed and acceleration data from "target" vehicles sampled from the population and 2) the use of onboard instrumentation in vehicles to record speed and acceleration data. Chase car data have primarily been used for developing driving cycles, while data from instrumented vehicles have been used only minimally. The use of other technologies, such as Global Positioning Systems (GPS), for collecting driving behavior data remains some time away from wide-scale implementation.

Given the importance of driving cycles to estimating mobile emissions, it is worth examining how the data are collected and the representativeness of the data for driving cycle development. Accordingly, this study has three objectives: 1) to assess the robustness of chase car data at a much finer resolution than previously examined, 2) to evaluate the appropriateness of mixing chase and target car data to develop the so-called composite driving cycles, and 3) to evaluate and recommend changes to minimize potential cycle construction biases that can arise as a result of chase car data collection procedures.

BACKGROUND

It is well established that the Federal Test Procedure (FTP), the foundation used for estimating mobile-source emissions inventories, does not adequately reflect normal driving patterns. Although still used in EPA-developed cycles, in 1990 the California Air Resources Board (CARB) initiated a project to develop new driving cycles to better represent actual driving behavior, thus improving mobile source emissions modeling (Gammariello and Long 1996). As part of this effort, driving data, specifically speed-time profiles, were collected on roadway networks in the Greater Metropolitan Los Angeles area in April and May of 1992. The resulting database is known as LA92, and data collection was accomplished using a chase car protocol.[1]

The LA92 chase car protocol was a refined version of procedures previously developed by General Motors (GM) and the U.S. Environmental Protection Agency (EPA). GM's approach involved a chase car following a vehicle from trip beginning to trip end and attempting to mirror the target vehicle's major speed changes, accelerations, and decelerations (Austin et al. 1993). The GM chase car was equipped with instrumentation for recording its own operations but not, however, with technology that allowed the recording of accurate estimates of target vehicle operations. Instead, the accuracy of the data hinged on the ability of GM chase car drivers to correctly match the speed and acceleration of target vehicles. In addition, the method was limited in that it did not account for the effects of changing road grades (Austin et al. 1993). The EPA protocol and equipment were similar to that of GM and also produced relatively imprecise speed-time profiles.

Two primary concerns arose with respect to this method. First, the potential for detection and resultant behavioral change by the target car driver was considered problematic. Second, the crudeness of acceleration and deceleration event measures likely resulted in inaccuracies in the data (Sierra Research 1997). To improve the resolution of previous measurements, the chase car used to collect data in the LA92 study was equipped with a range-finder laser designed to measure the relative distance between the chase car and target vehicles. A video camera, mounted inside the chase car, recorded the view through the windshield to provide a visual check for assessing data reliability. Presumably, the LA92 database more accurately captures the behavior of drivers in Los Angeles.

LA92 has been used by both CARB and EPA to create various driving cycles. The new EPA-constructed facility-specific cycles were developed

[1] A chase car protocol generally refers to the procedures used to identify a target vehicle and initiate speed-time data collection.

using the LA92 data in combination with the 1992 chase car data gathered in Spokane, Washington (S92) and in Baltimore, Maryland (B92), while CARB has used the LA92 data to develop the Unified Cycle (UC). The EPA speed-based cycles use speed, acceleration frequency, trip length, and level of service (LOS) variables to define cycles. The UC is a trip-based cycle that incorporates average speed, acceleration frequency, and trip length variables but does not include level of service. CARB has also constructed Unified Correction Cycles (UCCs) using LA92 to adjust for speeds between 10 and 50 mph. Since LA92 contains limited data at higher and lower speed ranges, CARB supplemented it with data from a separate 1992 EPA database that contained additional data collected in Baltimore, Spokane, and Atlanta using vehicles with onboard instrumentation. The EPA-instrumented vehicle data were used to supplement data for speeds between 0 and 10 mph and 55 and 75 mph.[2]

CURRENT STUDY METHODOLOGY

To evaluate the robustness of the chase car data and the appropriateness of mixing chase and target data, the present study considers three areas of concern: 1) potential inaccuracies in the data introduced by the current chase car protocol and equipment; 2) variation in the amount of data collected in Baltimore, Los Angeles, and Spokane and the data's representativeness of each region's traffic conditions; and 3) differences in driving behavior data recorded from target vehicles and from the chase vehicle when no targets are available, in other words, under "non-lock" conditions.

Potential Inaccuracies in the Data Due to the Current Chase Car Protocol and Equipment

Briefly, the current protocol, developed for the LA92 study, directs chase car drivers to collect second-by-second speed-time profiles from hundreds of target vehicles. The chase car follows predefined routes, "locking on" to target vehicles with the range-finder laser while simultaneously collecting data on such variables as road grade, type of vehicle targeted, road facility type, and level of service in addition to speed and acceleration (Austin et al. 1993). A full description of the chase car protocol can be found in Austin et al. 1993.

The way the chase car protocol is actually implemented during data collection can substantially affect the development of driving cycles. There are several potential problems that can cause the application of the protocol to vary. For the purposes of this study, there are two critically important instructions in the current chase car protocol: 1) the procedure for target car selection on busy surface streets and freeways and 2) the procedure for data collection under non-lock conditions.

The first procedure's objective is to ensure random vehicle selection. When chase car drivers enter a new roadway, the procedure instructs them to follow the first forward vehicle encountered in the same lane as the closest white vehicle. Specifically, the closest white vehicle is defined as "the closest white vehicle in front of an imaginary line passing through the center of the chase car and perpendicular to the direction of travel" (Austin et al. 1993, 53). If the chase car is in the same lane as a white vehicle and more than one white vehicle is present, "one car length is subtracted per 10 mph of speed before deciding which white vehicle is the closest" (Austin et al. 1993, 53).

The field application of this selection process can be complex under rapidly changing traffic conditions, making its execution very difficult. Review of videotapes recorded during the LA92 data collection indicated the procedure for selecting target vehicles was inconsistently applied, particularly when the chase car entered a new roadway. Many targets were not acquired even though the video suggested it was possible to do so. This appeared to be in part due to confusion about which vehicle should be chosen according to the target vehicle selection procedure.

With respect to the second procedure, chase car drivers are told to "drive in a fashion that approximately matches the general flow of through traffic," driving "faster than some vehicles and slower than a similar number of vehicles" in the absence of target vehicles (Austin et al. 1993, 54). In this

[2] It should be noted that although the LA92, B92, and S92 data sets were collected using the same protocol, to date only LA92 has been used on its own for driving cycle development.

case, the chase car records its own operating data with the range-finder laser disengaged. These data are then used to replace missing target data in the final "composite" database.

This use of chase car data in lieu of missing target car data is intended to increase the sample size available for building driving cycles. Target data and non-lock chase car data are joined together in series to create the composite data set. These composite data are then used to create Speed Acceleration Frequency Distributions (SAFDs), the cornerstone of driving cycle development. Since approximately 47% of LA92, 58% of S92, and 63% of B92 come from the chase car operations rather than the target vehicle, the driving behavior of chase car drivers and their ability to approximate the speed and acceleration of other vehicles become very influential. This is particularly true in light traffic conditions when there may be few vehicles to emulate. It should be noted that both the target and non-lock chase car data are recorded on a second-by-second basis, providing hundreds of profiles (realizations) of the sampling unit, that is, the driver-vehicle. When chase car data are used in place of missing target data, it does not increase the overall sample size of the data set. Instead, only one driver-vehicle profile, that of the chase driver, is added to the sample, increasing the sample size by one. Since this single profile will contain more speed-time data points than those of the target vehicles, there is great potential for the chase vehicle data to bias cycles developed from the data.

The choice of technology used for data collection can also have a significant impact on the successful application of the chase car protocol. The chase car has a built-in speed measurement system that records speed at every second with a precision of 0.38 mph (Austin et al. 1993). The range-finder laser installed behind the grill of the chase car emits 400 light pulses per second. When the laser beam bounces off the target vehicle, the time it takes for the signal to return to a receptor in the chase car's grill determines the distance between the two vehicles. The laser system was tested on static targets, yielding a distance accuracy within one foot. This presumably leads to a corresponding target vehicle speed error of 2 feet per second or 1.36 mph when the chase car is in motion (Austin et al. 1993). The potential errors, then, of the chase car speed measurement system and the range-finder laser together yield an error of ±1.74 mph in the estimated target vehicle speed.

This error is reasonable when estimating speed, but the impacts in terms of acceleration are less clear. If forward differencing is used to determine accelerations on a second-by-second basis (acceleration equals velocity at the second second minus velocity at the first second), the estimate of a target vehicle's acceleration could be off by as much as 3.48 mph/s.[3] While this represents the most extreme case, it illustrates how measurement errors could cause a target vehicle to appear to be accelerating when it is, in fact, at cruise.

Regional Differences in the Data

To examine data differences between cities, we analyzed the B92, S92, and LA92 composite data using two variables: level of service (LOS) and facility type. The B92 data include 191,119 seconds of data, representing 218 routes; S92 contain 175,137 seconds, encompassing 249 routes; and LA92 contain 102 chase car runs, resulting in 100,709 records (seconds) of data.

Variation in the Amount of Data Collected in Each Level of Service

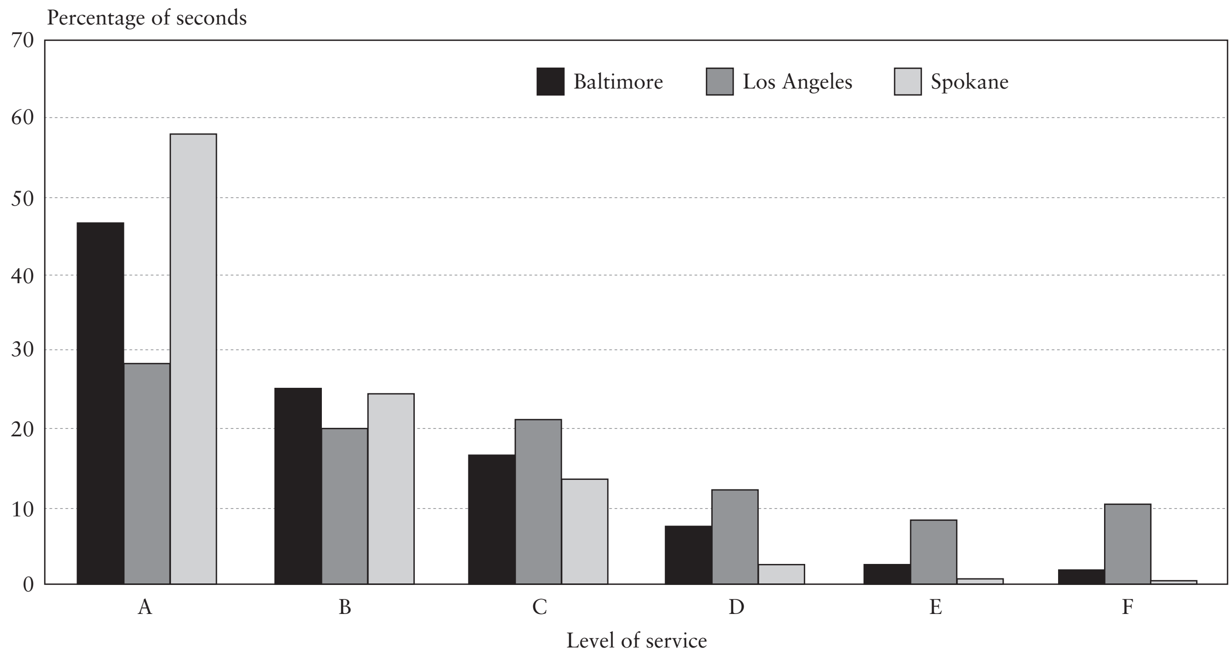
Level of service (LOS) refers to traffic density conditions observed on a specific facility at a specific time. In each of the chase car studies, the chase car observer visually assigned a level of service category (A, B, C, D, E, or F).[4] The observer used a switchbox mounted on the chase car's dashboard to manually record the level of service (Austin et al. 1993). Figure 1 presents the percentage of composite data by time collected at each LOS by city.

Despite differences in size, location, and availability of public transportation, Baltimore and Spokane exhibit similar amounts of time in levels

[3] Imagine that the velocity at the second second has an error of +1.74 mph, while the velocity at the first second has an error of −1.74 mph.

[4] These levels correspond to the levels of service given in the Highway Capacity Manual. Level A describes free-flow conditions, while Level F describes stop-and-go conditions. Levels B, C, D, and E represent levels of increasing traffic congestion.

FIGURE 1 Percentage of Seconds Spent in Each Level of Service (LOS) by City



of service B and C. Perhaps more interesting, however, is the difference in time spent at levels of service E and F between Baltimore and Los Angeles. The chase car recorded almost 3.5 times the amount of data in LOS E in Los Angeles as in Baltimore and approximately 6 times more data in LOS F. These numbers suggest substantial differences in traffic congestion level between cities.

One qualitative means to assess city to city difference in congestion is to use the Roadway Congestion Index (RCI) (Texas Transportation Institute 1998). The RCI is calculated as:

$$RCI = \frac{\frac{FwyVMT}{LnM} \times FwyVMT + \frac{ArtVMT}{LnM} \times ArtVMT}{13{,}000 \times FwyVMT + 5{,}000 \times ArtVMT} \quad (1)$$

where

- FwyVMT is the estimated vehicle-miles traveled on the chosen area's freeways.
- LnM is the estimated lane-miles of roadway.
- ArtVMT is the estimated vehicle-miles traveled on principal arterial streets.

The constants 13,000 and 5,000 indicate the capacity of the facility type, in this case freeway and arterial. The RCI is an indicator of average congestion over an entire metropolitan region and is used extensively by public officials.

According to the 1992 RCI, Los Angeles had a score of 1.54, while Baltimore had a score of 1.04 (Texas Transportation Institute 1998). Since chase car data should ideally characterize the driving behavior of the population and relative congestion levels within a metropolitan region, it is possible to use the ratio of different RCI scores as a basis for comparing the relative levels of service represented by the chase car data. On a relative basis, if driving data on congested roadways were sampled more often than indicated by the RCI ratios, the data would not be representative of average driving conditions in the area. Using this logic, LA92 should only contain about 1.5 times more records in levels of service E and F than B92. This suggests that there may be some unevenness in how level of service between cities was assigned or, alternatively, that different traffic levels were not appropriately sampled. Other studies have also suggested that the visually determined LOS may not represent level of service target statistics assigned by the Highway Capacity Manual's methods (Niemeier et al. 1998).

TABLE 1 Descriptive Statistics on Composite Speeds by Facility Type and Level of Service (LOS)

	Baltimore				Los Angeles				Spokane			
Fac./LOS	Min.	Max.	Mean	St. dev.	Min.	Max.	Mean	St. dev.	Min.	Max.	Mean	St. dev.
Arterial												
All LOSs	0	74.50	25.41	18.04	0	63.10	21.48	15.56	0	74.90	26.53	17.01
LOS A	0	74.50	30.97	17.77	0	61.10	22.65	15.21	0	74.90	29.25	16.92
LOS B	0	65.00	23.60	17.49	0	63.10	22.90	15.85	0	69.50	26.39	16.99
LOS C	0	63.15	20.30	16.60	0	59.10	21.24	15.65	0	56.68	18.95	14.35
LOS D	0	71.29	17.51	15.15	0	53.00	18.35	15.12	0	53.40	15.14	13.13
LOS E	0	56.77	12.15	14.77	0	47.69	13.39	13.74	0	34.50	15.21	11.22
LOS F	0	43.20	11.79	12.02	0	44.20	11.15	12.54	0	36.90	5.51	8.45
Ramp												
All LOSs	0	72.69	40.35	17.06	0	76.00	29.83	20.25	0	79.10	37.23	17.87
LOS A	0	69.80	42.50	15.49	0	61.10	28.88	19.16	0	79.10	37.67	18.79
LOS B	0	71.15	39.62	15.32	0	76.00	35.44	20.17	0	61.80	38.60	13.93
LOS C	0	68.52	42.52	15.61	0	65.70	30.42	20.96	0	63.70	33.28	19.72
LOS D	0	72.69	44.04	19.69	0	65.70	24.31	21.33	—	—	—	—
LOS E	15.31	47.60	28.50	9.69	0	63.70	38.59	16.05	—	—	—	—
LOS F	0	42.40	4.31	8.69	0	51.50	20.90	15.58	—	—	—	—
Freeway												
All LOSs	0	80.91	56.12	13.21	0	80.30	44.75	20.31	0	83.15	59.10	8.85
LOS A	0	80.91	59.39	7.07	29.60	76.36	57.27	8.72	0	83.15	62.06	9.86
LOS B	0	77.30	59.13	8.57	28.00	75.60	62.72	6.14	0	73.83	58.20	7.12
LOS C	0	75.84	59.24	8.34	13.72	80.30	60.85	7.86	0	70.70	57.00	7.19
LOS D	0	75.80	57.13	10.49	6.10	73.97	56.42	8.83	31.57	60.04	54.92	6.79
LOS E	0	71.30	44.37	16.33	0	69.50	38.89	17.42	24.95	55.70	41.72	6.34
LOS F	0	69.47	23.81	16.46	0	66.00	23.03	13.82	—	—	—	—

— Missing data

Differences Across Cities within Each Recorded Level of Service

Table 1 contains estimated standard errors and other descriptive statistics for all levels of service represented in the three databases across three main facility types (arterial/collectors, freeways, and ramps). From the table, it can be seen that the mean speed for level of service F on arterial/collectors is very similar for Baltimore and Los Angeles but differs substantially from that recorded in Spokane; that is, the composite data for Baltimore and Los Angeles show approximately twice the mean speed as Spokane and about 1.5 times the standard deviation. The ramp facility type shows even more pronounced differences between mean speeds in Baltimore and Los Angeles at all levels of service, particularly under more congested conditions. For example, the mean speed in Los Angeles' level of service F on ramps is nearly five times that recorded during Baltimore's level of service F on ramps. Further, the associated standard deviation for Los Angeles is almost twice the magnitude of that for Baltimore.

The RCI and the differences in mean speeds and standard deviations, however, suggest that there may be some unevenness in how level of service between cities was assigned or, alternatively, that different traffic levels were not appropriately sampled. The accurate reflection of level of service is particularly important because the EPA facility-based cycle depends on level of service as one of its key variables.

Target versus Non-Lock Chase Car Data on Each Facility Type

We also investigated the relative "lock-on" rates of chase to target vehicles with respect to different facility types and different levels of service. Lock-on rates indicate how much of the data in each database actually comes from target vehicles as opposed to chase car operations. When viewed by facility type, it is clear that lock-on rates vary between cities and, as such, might be expected to vary dramatically between facility types. Table 2 compares the percentages of data recorded from target vehicles with those recorded from the non-lock chase car in each city by facility type, using each city's composite data.

As previously noted, approximately 47% of the data in LA92 originates as non-lock chase car records, while approximately 58% of the Spokane data and 63% of Baltimore data come from non-lock chase car records. The implication is that the facility types with low lock-on rates are extremely dependent on the ability of a chase car driver to accurately mimic prevailing traffic conditions and/or to drive like the "average driver" if there are no other vehicles around. Lock-on rates are relatively low on private roads, local roads, and, to some extent, ramps and arterials/collectors, suggesting few targets or difficult terrain conditions. With few target vehicles on the road, the chase car driver necessarily has difficulty gauging how to drive "with the general flow," as described in Austin et al. (1993, 53).

Equally important is the time spent on each roadway type. For example, although Los Angeles high occupancy vehicle (HOV) lanes show 100% lock-on, when the overall time spent in HOV lanes is examined, the limitations of the collected data become apparent. Table 3 shows the percentage lock-on for LA92 according to the share of total time recorded by facility type.

In table 3, the time-weighted lock-on rates are the actual amount of target data as a share of the aggregate data. It is apparent that the largest combined share of time and lock-on occurs on arterial/collectors, 29.4%. It also is apparent that target vehicle data on ramps are very limited. Although target vehicle data make up 37.8% of the data recorded on this facility type, ramps represent a mere 5.4% of the seconds in the overall data set.

TABLE 2 Lock-On Rates and Non-Lock Chase Car Data by Facility Type Using Composite Data Sets

City	Facility type	"Lock-on" rate (percentage of data from target vehicles)	"Non-lock" rate (percentage of data from chase vehicle)	Total
Baltimore	Private road	4.2	95.8	100.0
	Local road	3.7	96.3	100.0
	Arterial/collector	38.3	61.7	100.0
	Ramp	23.9	76.1	100.0
	Freeway	61.8	38.2	100.0
	HOV lane	66.7	33.3	100.0
	Aggregate[a]	**37.2[b]**	**67.8**	**100.0**
Los Angeles	Private road	0.0	100.0	100.0
	Local road	1.5	98.5	100.0
	Arterial/collector	45.2	54.8	100.0
	Ramp	37.8	62.2	100.0
	Freeway	76.9	23.1	100.0
	HOV lane	100.0	0.0	100.0
	Aggregate[a]	**53.1**	**46.9**	**100.0**
Spokane	Private road	5.0	95.0	100.0
	Local road	5.3	94.7	100.0
	Arterial/collector	41.9	58.1	100.0
	Ramp	27.9	72.1	100.0
	Freeway	74.9	25.1	100.0
	HOV lane	—	—	—
	Aggregate[a]	**42.3**	**57.7**	**100.0**

[a] "Aggregate" refers to each city's composite data set without distinction between facility type.
[b] 1,458 seconds of data in the Baltimore database were undefined and, therefore, left out of this analysis.
— Missing data.

TABLE 3 Percentage of Target Data in Los Angeles as a Share of Total Time on Each Facility Type and in the Aggregate

Facility type	Lock-on rate (percentage of data from target vehicles)	Time on each facility type (percentage of total seconds)	Percentage of time-weighted lock-on rate
Private road	0.0	0.2	0.0
Local road	1.5	1.4	0.02
Arterial/collector	45.2	65.0	29.4
Ramp	37.8	5.4	2.0
Freeway	76.9	28.1	21.6
HOV lane	100.0	0.0001	0.0001
Overall rate	53.1	100.0	53.1

Multiplying 37.8% by 5.4% reveals that only 2.0% of the composite data were recorded from target vehicles on ramps. Since accelerations necessarily occur at these locations, ramps are important in defining mobile emissions. Target car data from these facilities would be much more useful for emissions purposes than non-lock chase car data.

Differences in Driving Behavior Between Drivers of Target and Chase Vehicles

To conduct side by side comparisons of the driving behavior of chase and target car drivers, we examined speed-time traces in the LA92, B92, and S92 data. Some interesting results emerged from this analysis. Two of these results are illustrated in the speed-time trace constructed from the LA92 data, figures 2a and 2b.

The thin line in the figure represents the non-lock chase car's speed-time trace, while the bold line represents the target vehicle's trace. At the point marked "A" in figure 2a, a target car abruptly appears in the data and is abruptly cut off, forming a hook shape. At the point marked "B" in figure 2b, the chase car accelerates just before acquiring the target and then begins to slow to match the target's speed.

We hypothesized that these peculiarities were caused from loss of vehicle lock and/or a chase car attempting to catch up to a prospective target. Review of the LA92 videotape revealed the following possible explanations for these anomalies.

- As hypothesized, suspect points, such as that marked "A" in figure 2a, are due to loss of lock on the target.
- Other suspect points, such as that marked "B" in figure 2b, are not due to the chase car accelerating to acquire a target. Instead, they appear to be primarily due to turning events and lane changes by target vehicles. In some cases, the chase car turns off one facility onto another. In other cases, the target vehicle simply exits the route from which the chase car is recording data.

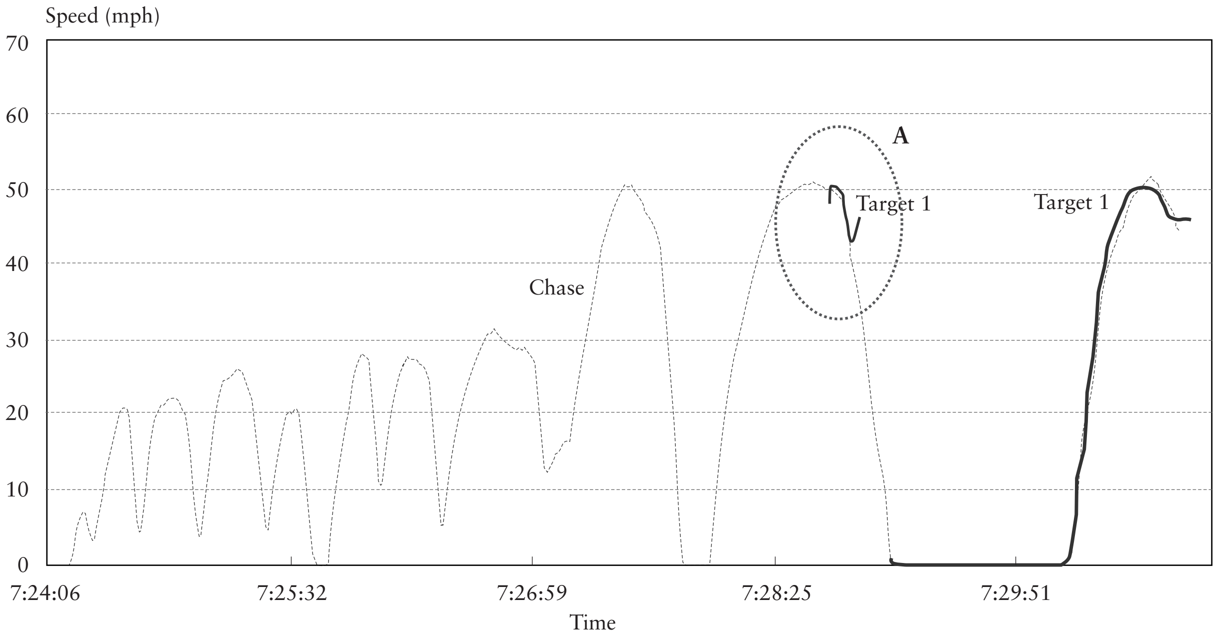
In addition to the LA92 videotapes, we reviewed videotapes pertaining to the 1997 Highway Performance Monitoring System (HPMS) data collection effort in the Sacramento, California region. The HPMS data were collected according to the LA92 chase car protocol and used the same chase vehicle as the LA92 project. The data exhibit similar anomalies, and a review of these additional videotapes corroborated our findings.

This analysis suggests that speed-time traces based on composite data are more representative of chase car operations than the target vehicle's. That is, chase car drivers do not seem to drive in a manner similar to the general public since they have different and specific motivations for their driving behavior. Consequently, the rationale for developing emissions cycles on the basis of these data may be invalid.

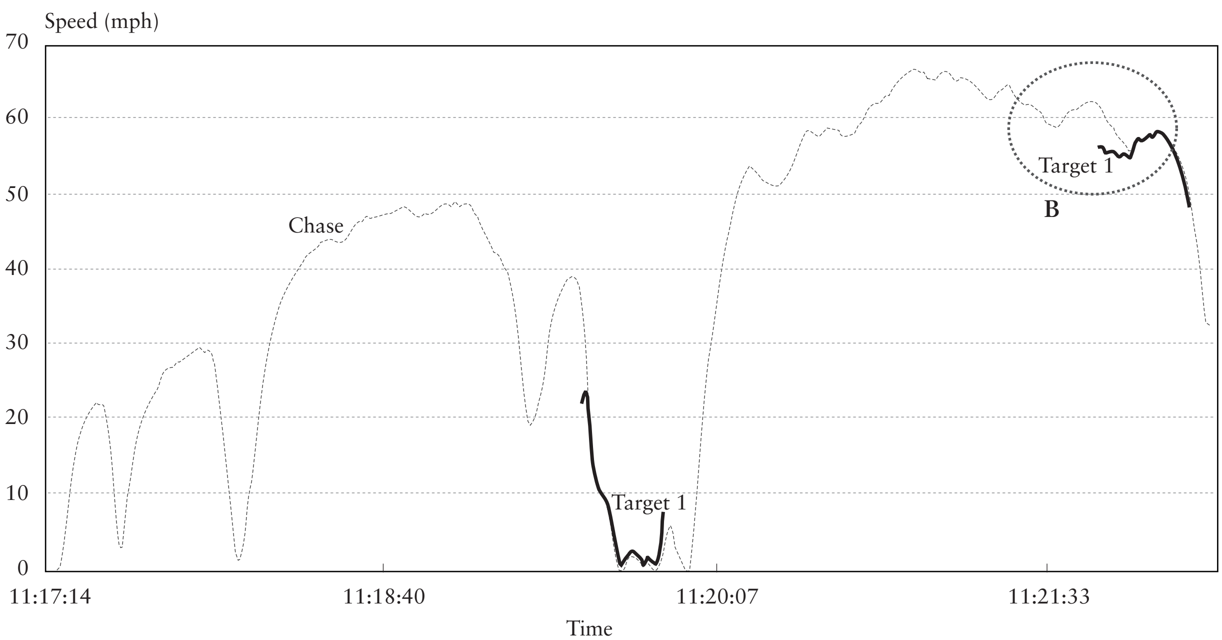
In order for non-lock chase car data to be representative of target vehicle data, the variation in speed under non-lock conditions should be equal to or less than the variation in target vehicle data. Although both lower and higher variation under non-lock conditions will weight the data disproportionately, smaller variation ensures a more conservative representation of driver behavior. A conservative perspective implies that chase car drivers tend toward the mean behavior of target vehicle drivers under similar conditions. The assumption underlying "conservative" is that the driving be-

FIGURE 2 Speed Time Traces of Chase and Target Vehicles from LA92

2(a) Speed time trace 04/07/92, run 1



2(b) Speed time trace 04/07/92, run 6



havior of the general population is represented by the mean speed of the target vehicles.

To assess if non-lock chase car data meet the conservative criterion, mean speeds and standard deviations of non-lock chase cars were compared to the mean speeds and standard deviations of target vehicles across all three cities. Since the percentages of lock-on vary widely by facility type and level of service, individual speed statistics were computed on the basis of these two factors. The HOV lane facility type has been removed from further consideration because there is so little data in this category. Table 4 contains mean speeds and standard deviations for both non-lock chase and

TABLE 4 Mean Speeds and Standard Deviations (mph) between Non-Lock Chase Car and Target Vehicle Data by Facility Type for All Three Cities

		Non-lock chase car (mph)		Target vehicle (mph)	
City	Facility type	Mean speed	St. dev.	Mean speed	St. dev.
Baltimore	All roadway types	28.2	19.7	35.7	22.0
	Private	11.4	10.9	14.3	12.0
	Local	18.2	13.0	17.1	15.5
	Arterial/collector	25.5	17.4	26.0	18.8
	Ramp	39.0	17.0	44.8	16.4
	Freeway	57.3	12.3	55.4	13.6
Los Angeles	All roadway types	26.6	18.5	30.0	21.4
	Private	5.6	4.7	—	—
	Local	16.4	11.1	16.4	11.8
	Arterial/collector	22.1	14.9	20.7	16.3
	Ramp	31.1	19.1	27.8	21.7
	Freeway	51.5	16.4	42.7	20.9
Spokane	All roadway types	27.7	18.2	32.7	20.8
	Private	5.3	6.8	1.3	2.6
	Local	21.1	17.1	37.4	15.5
	Arterial/collector	27.0	16.7	25.9	17.5
	Ramp	35.6	18.7	41.3	14.8
	Freeway	58.7	11.3	59.2	7.8

— Missing data.

target vehicles according to facility type and for all three cities.

In the aggregate data ("All roadway types"), B92 and S92 show larger differences in mean speeds between non-lock chase car and target vehicles than does LA92. In comparing mean speeds by facility type, we see that the most pronounced difference in Los Angeles occurs on freeways; in Baltimore, on ramps; and in Spokane, on local roads.

Table 5 shows the mean speeds and the results of a comparison of mean speeds by facility type using analysis of variance (ANOVA). The results of the ANOVA suggest that mean speeds of non-lock chase car and target vehicles, calculated on a second by second basis, are significantly different in all three cities and on nearly all facility types at a five percent significance level. The only facility type that did not show a significant difference in mean speeds was local roads in the LA92 database, but in this case the data were rather sparse.

An incremental difference in speed between non-lock chase cars and target vehicles is likely to be more important than identifying a large difference, particularly at higher speeds where coaxing the engine into an enrichment phase is likely to be accomplished with much smaller speed changes. To examine the speed variation from this perspective, the coefficient of variation (CV) of speeds was computed for target vehicles and for non-lock chase cars. Tables 6 and 7 contain these CV values for various facility types and levels of service represented in B92, S92, and LA92.

Table 6 contains the CV values for speeds computed on the basis of facility type. Observe that the target vehicle typically exhibits more variation in speed than does the non-lock chase car for LA92. As mentioned previously, to keep the estimates of driving behavior conservative, this should be the case. The coefficients of variation for LA92 suggest that the non-lock chase car data tend to reflect mean speeds, (i.e., the behavior of the "average driver,") rather than introducing additional variation into the data set. For B92 and S92, however, the results are less clear. For some facility types, such as local roads in Spokane and private roads in Baltimore, the variation in chase car speed far outweighs the variation in target vehicle speed.

TABLE 5 Mean Speeds and Comparison of Mean Speeds by Facility Type Using Analysis of Variance (ANOVA)

City	Facility type	Non-lock chase car mean speed (mph)	Target vehicle mean speed (mph)	Paired comparison of non-lock and target vehicle mean speeds (ANOVA) F-statistic	Significance
Baltimore	All roadway types	28.2	35.7	5,892.4	.000
	Private	11.4	14.3	6.1	.014
	Local	18.2	17.1	5.0	.026
	Arterial/collector	25.5	26.0	22.8	.000
	Ramp	39.0	44.8	157.1	.000
	Freeway	57.3	55.4	183.8	.000
Los Angeles	All roadway types	26.6	30.0	702.9	.000
	Private	5.6	—	—	—
	Local	16.4	16.4	.000	.998
	Arterial/collector	22.1	20.7	132.1	.000
	Ramp	31.1	27.8	32.9	.000
	Freeway	51.5	42.7	965.1	.000
Spokane	All roadway types	27.7	32.7	2,915.7	.000
	Private	5.3	1.3	26.4	.000
	Local	21.1	37.4	580.6	.000
	Arterial/collector	27.0	25.9	134.5	.000
	Ramp	35.6	41.3	62.957	.000
	Freeway	58.7	59.2	11.965	.001

— Missing data.

TABLE 6 Coefficients of Variation for Non-Lock Chase Car and Target Car Speeds in All Three Cities by Facility Type

	Non-lock chase CV	Target CV
Baltimore		
All roadway types	0.70	0.62
Private	0.95	0.84
Local	0.71	0.91
Arterial/collector	0.68	0.72
Ramp	0.44	0.37
Freeway	0.21	0.24
Los Angeles		
All roadway types	0.70	0.71
Private	0.80	—
Local	0.68	0.71
Arterial/collector	0.68	0.79
Ramp	0.62	0.78
Freeway	0.32	0.49
Spokane		
All roadway types	0.66	0.64
Private	1.30	2.05
Local	0.81	0.41
Arterial/collector	0.62	0.67
Ramp	0.52	0.36
Freeway	0.19	0.13

— Missing data.

In table 7, the CV values are disaggregated for the three predominant facility types (arterial/collector, freeway, and ramp) on the basis of level of service. As table 7 shows, speed is slightly more variable for non-lock chase cars than for target vehicles on freeways in Los Angeles under levels of service B, C, and D and on ramps under level of service C. However, table 7 indicates that in Baltimore there is significantly more variability in mean speeds when jointly considering level of service and facility type. The results in table 7 can be summarized by saying that chase car speeds exhibit greater variability under certain levels of service than do target vehicle speeds.

Accelerations and Decelerations

It is commonly known that acceleration events strongly affect emissions. However, Cernuschi et al. (1995) demonstrated that differences in relative deceleration rates have an effect on emissions as well. Consequently, in the present study acceleration and deceleration rates were compared for non-lock chase car and target vehicles represented in the LA92 database.

TABLE 7 Coefficients of Variation of Non-Lock Chase Car and Target Vehicle Speeds by Facility Type and Level of Service

	LOS A		LOS B		LOS C		LOS D		LOS E		LOS F	
	NL	TGT	NL	TGT	NL	TGT	NL	TGT	NL	TGT	NL	TGT
Baltimore												
Arterial	**0.57a**	0.54	**0.75**	0.68	0.78	0.81	**0.85**	0.84	1.08	1.28	**1.66**	0.74
Ramp	**0.38**	0.27	0.34	0.48	**0.43**	0.25	**0.51**	0.16	**0.34**	0.23	2.02	—
Freeway	**0.13**	0.11	0.14	0.14	**0.18**	0.11	**0.27**	0.14	**0.39**	0.36	**0.69**	0.68
Los Angeles												
Arterial	0.66	0.71	0.65	0.73	0.69	0.77	0.81	0.83	0.94	1.06	0.52	1.34
Ramp	0.66	0.70	0.55	0.59	**0.74**	0.60	0.60	1.24	0.34	0.51	0.60	0.87
Freeway	0.14	0.16	**0.11**	0.09	**0.14**	0.13	**0.19**	0.14	0.39	0.46	0.56	0.60
Spokane												
Arterial	**0.58**	0.57	**0.65**	0.64	0.76	0.76	0.77	0.89	**1.10**	0.69	**1.61**	1.22
Ramp	**0.51**	0.44	**0.46**	0.23	**0.66**	0.43	—	—	—	—	—	—
Freeway	**0.22**	0.13	**0.20**	0.08	0.12	0.13	**0.22**	0.12	0.12	0.15	—	—

NL = Non-lock chase car
TGT = Target vehicle
[a] **Bold numbers** indicate non-lock chase car CVs that are higher than the CVs of corresponding target vehicles.
— Missing data.

Three categories were created: cruise (–0.0340 ≤ a ≤ 0.0340 mph/s), normal acceleration (0.0341≤ a ≤ 3.290 mph/s), and hard acceleration (a ≥ 3.30 mph/s), where a is the second-by-second acceleration of the vehicle. The cruise interval is based on work by Holmén and Niemeier (1998), and the hard acceleration interval is based on previous evaluations of chase car data (Austin et al. 1993). Deceleration rates were classified as mirror images of their acceleration counterparts: cruise (–0.0340 ≤ a ≤ 0.0340 mph/s), normal deceleration (–0.0341≤ a ≤ –3.290 mph/s), and hard deceleration (a ≤ –3.30 mph/s). Table 8 indicates the percentage of time that chase cars and target vehicles spend in the various acceleration and deceleration intervals.

TABLE 8 Percentages of Time Spent in Accelerations and Decelerations by Chase and Target Vehicles (LA92)

	Chase (non-lock)	Target
Acceleration		
Cruise	27.9	21.4
Normal	66.3	75.2
Hard	5.9	3.4
Total	**100.0**	**100.0**
Deceleration		
Cruise	31.0	24.5
Normal	55.7	70.5
Hard	13.3	5.0
Total	**100.0**	**100.0**

Although the percentage of time that non-lock chase cars and target vehicles spend in the hard acceleration and hard deceleration intervals is small relative to the percentage of time spent in normal and cruise categories, the analysis indicates the differences in time spent in hard accelerations and hard decelerations are substantial. Non-lock chase cars recorded almost twice as many hard accelerations as did target vehicles and over 2.5 times as many normal acceleration events as target vehicles.

A relatively large number of the hard accelerations/decelerations may be explained by the chase car drivers' need to speed up or slow down to quickly acquire a target. However, since the non-lock chase car data is used to replace missing target vehicle data, the imputed values will be very influential and can contribute to an overprediction of modal frequency. If the overprediction is considerable, driving cycles will likewise tend to have too many of these modal events.

CONCLUSIONS AND RECOMMENDATIONS

Currently, composite chase car data provide the foundation from which driving cycles are developed, both to ensure vehicle compliance with air quality regulations and to characterize average emissions in the overall fleet. The robustness of the data can be judged by examining variation when the data are disaggregated by region, level of service, and facility type. The sources of this variation described and analyzed in this paper include difficult chase car protocol instructions, differences in the amount of data collected in each city, and differences between the driving behavior of target vehicle drivers and non-lock chase car drivers.

Two conclusions can be drawn.

- Instrument failures and misapplication of data collection procedures result in anomalies that significantly impact overall data representativeness.
- The composite data currently used for driving cycle development, which combine both non-lock chase car and target vehicle measurements, do not contain enough target vehicle information to adequately reflect the driving behavior of the general population.

These conclusions suggest that changes in the chase car protocol and in the technology used to measure target vehicle speed and acceleration could reduce biases in emissions cycles developed using the composite data. To address these problems in future chase car data collection efforts, we briefly elaborate on some recommendations.

Instrument Failures and Misapplication of Data Collection Procedures

Anomalies in speed-time traces from LA92 and HPMS data appear to be caused by protocol and/or instrument failure. Such anomalies may significantly influence cycles developed using the data sets and, therefore, the construction of driving cycles. The most notable deficiencies with respect to the existing technology are 1) the inability of the range-finder laser to maintain a lock on target vehicles when going over bumps, around slight curves, or on changing road grades and 2) the potentially large errors in measuring target vehicle accelerations. These can lead to both a marked lack of target data on ramps and inclines and a misrepresentation of target vehicle modal frequency. Currently, the most feasible and effective changes in technology would involve improvements to the laser or possibly the development of an appropriate scanning radar or scanning lidar system.

Problems with Mixing Non-Lock Chase Car and Target Vehicle Data

When examined at a fine scale of resolution by facility type and level of service, substantial variation, attributable to driving behavior, is observed in the composite data of B92, LA92, and S92. Significant differences in mean speeds between target and non-lock chase vehicles show that target drivers and chase car drivers represent separate populations. Similarly, an examination of accelerations and decelerations reveals disproportionate variation between non-lock chase car and target vehicle drivers. Given the uses of chase car data, it is important to note that combining the data may mask important differences between drivers in the resultant driving cycle. Therefore, we recommend that non-lock chase car data be minimized in, if not eliminated from, the driving cycle development process.

Proposed Changes in the Current Chase Car Protocol

Three changes in the chase car protocol can be made to create more robust databases at every level of aggregation. A primary element is the acquisition of additional target vehicle data. To acquire these data, the protocol should contain 1) simpler chase car routes and target car data collection procedures, 2) a simplified target vehicle selection procedure, and 3) the use of a traffic density measure rather than a visual assignment of level of service.

As evidenced by the LA92 videotapes, chase car routes appear to be too complicated for the drivers to concentrate on collecting target vehicle data. Various improvements to the route design could be made. One method would be to divide routes, predetermined from the top origin-destination pairs in a region, into segments by facility type and to have chase cars collect repeat data on the same segment in order to characterize target drivers' behaviors on that facility. The overall objective would be to allow chase car drivers to choose greater numbers of target vehicles and potentially stay with them for longer periods of time, thus increasing the records of target vehicles available in the databases. Simplifying the route design would also remove the chase car drivers' disincentive to engage target vehicles by leaving the chase car on a single facility for a longer period of time. The implicit assumption in this approach is that driving behavior on like facilities is similar.

Although the current vehicle selection method randomizes target vehicles and captures lane variation on multilane facilities, a revised lane sampling program based on predetermined lane choices would be less complex and would result in more reliable target car data. Any change in the lane-sampling program should be implemented in such a way as to guarantee all sources of variation are adequately represented. Two specific sources are within-lane variation and between-lane variation. Within-lane variation encompasses differences among potential target vehicles driving in the same lane, and sampling more vehicles and a wider range of vehicle types can adequately represent it. Capturing between-lane variation requires more extensive pre-run planning, especially on freeways where the sampling bias appears to be most extensive.

Finally, the collection of visually assigned level of service measurements appears to be of questionable use. As reported here, determination of level of service during chase car runs is very inconsistent. However, our analysis of data on the basis of level of service suggests it significantly affects mean speeds on certain facility types across all three cities. Poor level of service determinations may be the result of the data recording procedure rather than a reflection of actual driving behavior. Given its subjective basis, visually assigned level of service is not a reliable parameter for use in construction of driving cycles. Density measures, perhaps compiled from local travel management centers, would be more appropriate for use in regional driving cycle construction.

ACKNOWLEDGMENT

The authors gratefully acknowledge the insightful comments and suggestions of the reviewers and the issue editor. This study was funded under the CAMP Project by the California Department of Transportation. The views reflect those of the authors alone.

REFERENCES

Austin, T., F.J. DiGenova, T.R. Carlson, R.W. Joy, K.A. Gianolini, and J.M. Lee. 1993. *Characterization of Driving Patterns and Emissions from Light-Duty Vehicles in California, The Final Report.* Sacramento, California: Air Research Board, A932-185.

Cernuschi, S., M. Giugliano, A. Cemin, and I. Giovannini. 1995. Modal Analysis of Vehicle Emission Factors. *The Science of the Total Environment* 169, nos. 1–3: 175–83.

Gammariello, R.T. and J.R. Long. Development of Unified Correction Cycles. Paper presented at the sixth CRC On-Road Vehicle Emissions Workshop, San Diego, CA, March 1996.

Highway Capacity Manual. 1994. Washington, DC: Transportation Research Board, National Research Council.

Holmén, B.A. and D.A. Niemeier. 1998. Characterizing the Effects of Driver Variability on Real-World Vehicle Emissions. *Transportation Research Part D* 3, no. 2: 117–28.

Niemeier, D.A., J.M. Utts, K. Nanzetta, and L. Ring 1998. *Speed Correction Curves: Scoping Study.* Institute of Transportation Studies, University of California, Davis, UCD-ITS-RR-98-12.

Sierra Research Inc. 1997. Average Speed Measurement in the SACOG Region: Technical Appendix. Report No. SR 97-09-01, California Department of Transportation.

Texas Transportation Institute. 1998. *Urban Roadway Congestion, Annual Report.* Available at http://mobility.tamu.edu/study/rci.stm.