The objective of this study is to estimate the fractions of various levels of high evaporative emissions across the Denver fleet. Approach used in the study relates indices of evaporative emissions (EI23) that were derived from RSD readings to levels of PSHED hot-soak evaporative emissions based upon correlations of a smaller stratified set of paired PSHED and RSD readings. The experimental approach and methods are adequately documented in the report and accompanying background document. Presentation of the results, including tables and figures, are generally clear except as noted in the following comments.

The data show that the measured PSHED 15 minute hot-soaks emissions are correlated with EI23, but with considerable scatter (Figure 4-1). We can see from this plot that the detection limit for the RSD EI23 index is poor and considerably worse than for the SHED measurements. Most of the EI23 values are clustered around 100 with corresponding PSHED emissions ranging from 0.01 to 20 g/Qhr. The preliminary study with induced evaporative emissions showed that the RSD evaporative index had a 50% chance of detecting evaporative emissions with PSHED-equivalent running loss level corresponding to about 20 g/Qhr (equivalent to EI23 Bins of 5 or below). While the EI23 evaporative index would be useful for identifying gross evaporative HC emitters, its ability to estimate fractions of high evaporative emissions within various levels of evaporative emission other than the top end of the distribution seems limited.

Conversion of EI23 measurements to Bins provides what appears to be clearer summary of the distribution of EI23 values by PSHED-equivalent running loss levels. As I understand this procedure, this classification assigns the estimated evaporative indices into bins with width that each corresponds to one standard deviation of the variability of a single EI23 measurement (after accounting for the effects of the exhaust HC emissions on EI23). The EI23 Bins are then associated with probabilities of exceeding various threshold PSHED hot-soak emission levels. This approach allows the association to be made without regard to the quality of the correlation between EI23 and PSHED hot-soak levels, which we know is poor. EI23 values in at least the first three EI23 Bins (with PSHED thresholds of greater than 1, 2 and 5 g/Qhr) are probably below the method limit of detection and are really random noise. If so, there is about equal chance that any of the EI23 values in the first three Bins has a corresponding PSHED above the threshold. Therefore, it is not unexpected that fractions of elevated PSHED in Table 4-6 are about the same for Bins 1 (6.7%), 2 (7.6%) and 3 (9.6%). These fractions are likely not valid given the measurement sensitivity. If 20g/Qhr is a reasonable level where the corresponding EI23 values become reliable, then the distribution shown in Table 4-4 for this High PSHED definition is valid for all EI23 Bins. The fractions are progressive less reliable for the lower EI23 Bins at lower thresholds values.

I believe the net result is an overestimation of the fractions of elevated PSHEDS in the lower Bins. Products of these fractions with the proportionally larger numbers of vehicles in these bins for the Random fleet will result in larger fractions of elevated PSHEDs in the larger fleet of vehicles. For example, results of the de-stratification calculations in Table 4-6 shows that 12.7% of the 5830 vehicles in the random sample are estimated to have corresponding high-PSHEDs defined as greater than 2 g/Qhr. If the first three Bins are counted as zero, then this fraction drops to 5.5%. Also dropping Bins 4 and both 4 and 5 reduces the fraction to 2.9% and 1.6%, respectively. The more appropriate fraction is likely between 1.6 to 5.5% rather than 12.7%.

It should also be noted that the distributions are presented without quantitative estimate of uncertainty and bias that are inherent in the study approach. In addition to the poor limits of detection of RSD evaporative index, the following sources of uncertainty and bias were not assessed in the report.

- The distributions are based on static SHED 15-minute hot-soaks and do not include diurnal evaporative emissions and may not fully account for all running emissions.

- The residual hydrocarbon signal in the RSD measurements in excess of the regression line of HC with CO_2 results is a crude measure of the diluted mixture of evaporative emissions from fuel permeation, vaporize fuel leaks, and fuel system venting during vehicle operation. Unlike exhaust pollutant, there are no tracers for evaporative HC emissions to account for dispersion rate of emissions.

- Replicate LSHED and PSHED tests have large variability. Section 4.5 does not address the significance of the large variability of replicate SHED tests to distribution of fractions of "high evaps" at various definitions.

Other General Comments

1. Ambient temperature was not included as a variable in the study design and PSHED and replicate RSD measurements were all made within a short time at about the same temperature. The test sets within each EI23 Bin were conducted at ambient temperature spanning a range of up to about 30 ^{o}C. Evaporative emissions are known to increase with ambient temperature with doubling of permeation for 10 ^{o}C rise in temperature. This likely would not be issue if ambient temperature was a random variable in the study and test sets within each bin had similar random distribution of temperature. Was this checked? The potential bias due to differences in temperature would be minimal for the high emitter bins, but may be more important for the other bins.

2. It would be helpful in Section 2 (Background) to state how the results of this study and similar future studies will be used in the MOVES model. Should be specific enough to identify the relevant algorithms and inputs.

3. Most vehicles in Bins 6 and 7 had high exhaust HC emissions, which can contribute to the estimated evaporative emissions. The report asserts that this positive interference is mitigated by the binning procedure. From the relevant discussion in Appendix A, it is difficult to determine the significant of the positive interference or the effective of the binning procedure.

4. The report does not include a summary of other testing – modified California Method (olfactory, visual and electronic HC sniffer examination of various vehicle components). If this information is summarized elsewhere, it should be references and a brief summary of the finding should be included within this report.

Specific Comments

1. P. 1-2, line 5. Are there plans for follow-on uncertainty analysis that can be described here?

2. P. 2-2, second full paragraph: Describe briefly the evidence, with appropriate references, that previous estimate of "high evaps" were lower that what is occurring in the real world.

3. P. 3-14, last sentence: Meaning is unclear. Why would large variability of PSHED hot-soaks itself result in overestimation of fraction of vehicles with high hot-soak emissions?

4. P. 4-1, 2nd paragraph, line 13: Rather than "accuracy", "representativeness" may be more appropriate in this context.

5. P. 4-3, 1st paragraph, last two sentences: States that influence of variability of hot-soak emissions will be discussed later in the section. This discussion appears to be missing.

6. P. 4-3, 2nd full paragraph: References to "not simulated exhaust" and "natural exhaust" in the last two sentences are confusing.

7. P.4-4, 1st full paragraph, last two sentences: The reason for selecting Method B is difficult to understand without prior knowledge that EI23 is based on residuals of the linear regression. This is only explained in Appendix A. It should be mentioned briefly in the Section 4.2 for clarity.

8. P.4-11, Table 4-2. VDF is not defined anywhere in the report.

9. P. 4-24, 1st paragraph, last sentence: Are the quantifications of uncertainties and bias part of a follow-up report? When is this expected?

10. P. 4-25, Table 4-6: What is the basis for S_h in the calculation of standard error of the fraction of elevated PSHEDs? What are the sources of the values used in calculating the standard deviation?

11. P. 4-30, Table 4-10. Unless there is good reason for using natural log, give estimated error for column 2 in units of g/Qhr.

12. P. A-1, item i): Residual rather than N?

13. P. A-2: Add a description of the origin of the constants used in equations shown at the bottom of the page. Explain how this reduces dependence of EI23 on exhaust HC concentrations.