888-242-9714 525-3018

Wert, Mark (DEP)

From: Richard Joy [RJoy@SierraResearch.com]
Sent: Wednesday, May 16, 2001 8:05 PM
To: Mark Wert (E-mail)
Subject: URGENT - Preliminary MA31-to-IM240 Correlation Results


excess.lst

Per your request, attached is a text file whose top section contains the results of Sierra's preliminary analysis of the MA31-to-IM240 correlation results. These results are based on our analysis of valid test results from roughly 275 vehicles that have undergone correlation testing. Note that the summary results shown in the file are based on simple averages of all AZ vehicles included in the initial data set; i.e., no attempt was made to weight the data to accurately reflect the model year and vehicle type distributions of the MA fleet. Also, the AZ vehicles were intentionally selected using a stratified sampling protocol designed to collect of data from a disproportionate fraction of high emitting vehicles. This approach was used to ensure that the resulting correlations would be representative of the entire range of vehicle emissions scores found in the MA fleet.

Given the above cautions, the results can be summarized as follows:

* Total excess emissions identified:
  - HC = 96%
  - CO = 98%
  - NO = 98%

* Total test vehicles = 275
* Total MA31 failures (based on current standards) = 84
* MA31 failure rate = 84/275 = 30.5%
* Total IM240 failures (based on EPA startup standards) = 43
* IM240 failure rate = 43/275 = 15.6%
* Total MA31 false failures (relative to the IM240 results) = 46
* Total MA31 false fail rate = 46/84 = 55%

These results are of substantial concern. The false failure rate of 55% is the highest by far ever seen by Sierra. To put this in perspective, most sets of cutpoints are designed to limit the false failure rate to no more than about 5%. The fraction of excess emissions identified relative to the IM240 is very high, but this is deceptive. It is largely a result of an overall MA31 failure rate that is roughly double that of the IM240. The MA31 failure rate is also much higher than the 10-12% emissions failure rate currently being seen in the MA program, based on the test results recently provided by Deems. A large part of this is likely due to the two issues (nonweighting of the AZ results and the testing of a disproportionate share of high emitters) identified above. However, IM240 failure rates are only a little higher than those actually being seen in the AZ program, which means the fraction of high emitters is not that far out of line with the actual AZ fleet. The relative relationship between true and false MA31 failures (which are roughly equal) is also a big concern. This means if you loosen the cutpoints to drop the false failure rate you are also likely to drop the true failure rate significantly, which will clearly have a negative impact on the excess emissions identification rate.

Another obvious question is why the failure rate seen in the correlation data is so much higher than the current fleetwide failure rate in MA. Even if the correlation data failure rate is discounted significantly to approximate the possible effect of the issues of fleet weighting and the

high emitter fraction on these results, it appears that the MA failure rate is considerably lower than it should be. Possible additional causes include:

* ESP equipment problems that are reducing the fail rate relative to what you would expect based on the AZ results collected on an SPX workstation. However, Deems recent email indicated the SPX and ESP failure rates were within 0.5% of one another so this appears to be a non-issue. (Deems didn't indicate if these were overall failure rates or emissions-only failure rates, so this is still a little bit of an open issue.)

* Unknown differences in equipment performance between the SPX workstation being used in AZ and the typical workstation located in the MA stations. The AZ workstation has obviously been fine-tuned to improve its performance; however, this effect should be fairly minor given that it is a standard production unit that is running the same software as contained on the MA workstations.

* A very high degree of improper testing is occurring in MA. The easiest of these would be for inspectors to be pulling the sample cone back from vehicles when they are being tested. Fuel economy checks are being conducted on every test but the software element that would reject tests that show too high of fuel economy results is not yet active. This means there is no fuel economy or exhaust flow check that would catch inspectors who are cheating in this manner. There are also many other ways (e.g., clean piping) they could be cheating.

We re not certain that a lot of fraudulent testing is occurring, but there do not appear to be many other possible explanations for the difference between the MA and AZ results. It is clear that a priority needs to be put on trying to better understand what is going on. This includes the following elements:

* Completing the correlation analysis as outlined in Sierra's memo of 3/5/2001. This includes normalizing the AZ results to the MA31 emissions rates to determine MA fleet-based excess emissions identification rates.

* The above element will NOT address the issue of the difference in MA versus AZ failure rates. There is no exact way to do this. However, the AZ results can be normalized by model year and vehicle type to the MA fleet, thus eliminating possible bias due to this factor. A closer evaluation of the results can then be made, including comparing the normalized failure rates and average emissions scores by pollutant, model year and vehicle type. It is unclear at this point how much insight this approach will yield; however, given the severity of the concern I suggest that we do this. It will not take long to do and may show some results that could shed some light on what is going on.

* An open question is if it is possible to somehow normalize the fraction of high emitters to the fraction that would be expected in each vehicle type/model year range grouping. I'm unsure if this can be done but we will talk internally here at Sierra regarding whether this is possible.

There may be some additional suggestions that Garrett, Michael and Phil have about what more can be done to figure out the causes of the huge difference in AZ and MA failure rates. I'll talk to them and let you know if we come up with any other ideas. A key issue is whether you want to wait for George to complete the correlation analysis or have us complete it independent of him to give you faster results. We can perform the second element described above (if we are provided with MA test results showing model year and vehicle type distributions) independent of the correlation study; however, finishing that analysis may provide some additional insight into what is going on.

The most critical issue as I see it is trying to better figure out whether a

higher share of high emitters in the AZ data is responsible for a lot of the difference in MA versus AZ failure rates. As I pointed out above, this doesn't seem to be based on the IM240 failure rates seen in the AZ test data; however, additional focus on this issue is clearly needed. It is also noted that these results are based on a little more than one-quarter of the total number of test results that is supposed to be collected in the correlation study. These results may therefore change somewhat as more data are collected. However, I do not expect very significant changes other than may result from a change in the fraction of high emitters included in the test data.

<<excess.lst>>

Richard