

Determination of PEMS Measurement Allowances for Gaseous Emissions Regulated Under the Heavy-Duty Diesel Engine In-Use Testing Program

Part 3 – Results and Validation

Christopher A. Sharp and Michael D. Feist
Southwest Research Institute® (SwRI®)

Christopher A. Laroo and Matthew W. Spears
U.S. Environmental Protection Agency

Copyright © 2009 SAE International

ABSTRACT

Beginning in 2007, heavy-duty engine manufacturers in the U.S. have been responsible for verifying the compliance on in-use vehicles with Not-to-Exceed (NTE) standards under the Heavy-Duty In-Use Testing Program (HDIUT). This in-use testing is conducted using Portable Emission Measurement Systems (PEMS) which are installed on the vehicles to measure emissions during real-world operation. A key component of the HDIUT program is the generation of measurement allowances which account for the relative accuracy of PEMS as compared to more conventional, laboratory based measurement techniques. A program to determine these measurement allowances for gaseous emissions was jointly funded by the U.S. Environmental Protection Agency (EPA), the California Air Resources Board (CARB), and various member companies of the Engine Manufacturer's Association (EMA). This program was conducted at Southwest Research Institute® (SwRI®), and was directed by a Steering Committee comprised of representatives from all of the funding parties. In field validation testing was conducted by the Center for Environmental Research &

As part of a three paper series, this paper will focus on the final results of program. The results of the statistical model used in the program will be presented, as well as the model validation efforts, and the analysis of the

results leading to the final measurement allowance values. Earlier papers focused on the laboratory test methods and on the details of the statistical methods.

INTRODUCTION

BACKGROUND

Based on a Settlement Agreement dated June 3, 2003, EPA and EMA agreed to a regulatory proposal for a manufacturer run Heavy-Duty In-Use Testing program. As stated in a subsequent EPA Regulatory Announcement dated June 2004, "Under this in-use testing program, manufacturers will measure gaseous and particulate exhaust emissions from diesel engines using portable onboard emission measurement systems. This cooperative effort represents a significant advance in helping to ensure that the benefits of more stringent emission standards are realized under real-world driving conditions"[1], [2], [3]. The final rule for the HDIUT program was announced in June 2005 [4], [5].

A Memorandum of Agreement (MOA) dated May 2005 defined a cooperative effort between EPA, EMA, and CARB "with regard to developing data-driven emission measurement accuracy margins for gaseous and PM emissions to be applied to the results of PEMS in-use in field testing under the forthcoming HDIUT regulations" [6]. Furthermore, a Test Plan dated May 20, 2005 was

developed by EPA, EMA, and CARB to define procedures and methodologies used to determine the gaseous measurement allowances [7].

This effort was guided in part by experiences gained during an earlier “pre-pilot” program which was run at CE-CERT comparing laboratory reference and PEMS measurements [8]. Evaluation of PEMS versus laboratory equipment has also been carried out on a light-duty chassis dynamometer [9].

OBJECTIVE

The HDIUT gaseous measurement allowance program was performed to experimentally determine additive measurement accuracy margins for in-use compliance testing using PEMS. The first paper in this series gave an overview of the program, and provided details regarding the individual experiments that were run to investigate various sources of PEMS measurement error [10]. The second paper described the statistical methods used during the program, and described the model simulation that was developed [11]. This third paper focuses on the model results, validation exercises, and analysis. Comprehensive final reports were prepared for EPA that contain details of the program [12], [13], while the measurement allowances generated during the program were issued in a direct final rule in February 2008 [14], [15].

PROGRAM OVERVIEW

A full description of the engines, vehicles, experimental methods, and analytical techniques used in the program is given in the other papers in the set [10], [11]. As an aid to the reader, a brief overview is given here.

The gaseous pollutants examined in the program were carbon monoxide (CO), non-methane hydrocarbons (NMHC), and oxides of nitrogen (NO_x). The program focused on development of a statistical model, using Monte Carlo techniques, which was developed to incorporate PEMS measurement errors from a variety of sources. The model was iterated over a wide variety of NTE data sets millions of times, to simulate various combinations of PEMS measurement errors. The model results were used to determine additive brake-specific measurement allowances for each pollutant.

The engine laboratory experiments used to generate data for the model were run using three engines, all of which were 2004 model year engines. These were a DDC Series 60 14L engine, a Caterpillar C9 engine, and an International 6L engine. All three engines were retrofitted with a catalyzed DPF for this program.

All of the PEMS were SEMTECH-DS instruments, supplied by Sensor Inc., and most of the engine measurements were conducted using three PEMS instruments in parallel at any given time. A total of seven SEMTECH-DS units were used over the course of the program. It had been originally planned that several

Horiba PEMS would also be tested, but these were not available at the start of the program. All of the model experiments were run at Southwest Research Institute. The on-road measurement data used to validate the model were generated by CE-CERT using their Mobil Emission Laboratory facility.

The entire program was conducted under the direction and guidance of a Steering Committee which was made up of representatives from EPA, CARB, and various EMA member companies.

MODEL RESULTS

The complete model run involved a total of 195 reference NTE events, each of which was run through the model at least 10,000 times. The final output for each individual model run was a “delta” or difference between the calculated pollutant level with all errors applied and the “ideal” calculated level (e.g. with no errors applied) for the given reference NTE event. The 10,000 (or more) trials conducted on each reference NTE event produced a distribution of these deltas, which was then evaluated to determine a 95th percentile delta for each pollutant, and each of the three calculation methods. This set of 95th percentile deltas formed the final output result of the statistical model. The three calculation methods for emissions are fully described in the second paper in this series; however, the calculation formulas are also included in this paper in the Appendix for reference as an aid to interpreting the results.

As described in the second paper of this series, the results of the Full Model run included all error surfaces which were incorporated in the model, and the time alignment adjustment. Simultaneously, a Validation Model run was conducted, which generated a set of error distributions for each reference NTE event wherein some of the error surfaces were turned off. The results of this Validation Model were used for comparison to the on-road tests conducted by CE-CERT in order to validate the model. This validation process and results are described later.

NO_x EMISSIONS

The results of the Full Model run for BSNO_x are shown in Figure 1. Each data point represents the 95th percentile delta for a single reference NTE event. These results are plotted against the “ideal” calculated NO_x value for the given reference NTE event. The collected results are shown for all three calculation methods. The NTE threshold is also shown for reference at 2.68 g/kW-hr.

As can be seen in Figure 1, all of the 95th percentile deltas were positive, indicating a positive error over the ideal reference value. Method 1 showed a significantly larger positive error than the other two calculation methods. Method 3 demonstrated a somewhat larger positive trend than Method 2.

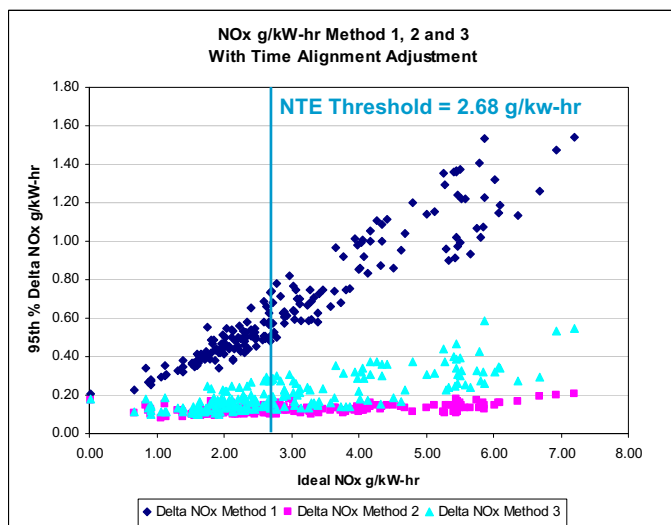


Figure 1. Full Model Results for NO_x

As mentioned in the second paper, additional analysis was conducted in order to determine which error surfaces were the most significant contributors to both variability and bias for each calculation method. A summary of this analysis for NO_x is shown in Figure 2. Where bias effects were observed, the direction of the bias is indicated by a (+) for positive offsets or a (-) for negative offsets. In the table OEM Torque refers to error data provided by engine manufacturer's intended to account for errors in the ECM broadcast torque level. DOE BSFC refers to a "design of experiment" data set generated in the program to account for the effect of ambient conditions (altitude, temperature, etc.) on ECM broadcast BSFC.

Variation (Spread)	Bias (Offset)
Method 1	Method 1
<u>Steady-State NOx</u>	<u>Steady-State Exhaust Flow (+)</u>
<u>Warmup Torque</u>	<u>Exhaust Flow Pulsation (+)</u>
OEM Torque	
Steady-State Exhaust Flow	
Method 2	Method 2
<u>Steady-State NOx</u>	<u>Steady-State CO₂ (-)</u>
<u>OEM BSFC</u>	<u>DOE BSFC (-)</u>
DOE BSFC	
Warmup BSFC	
Method 3	Method 3
<u>Steady-State NOx</u>	<u>Steady-State CO₂ (-)</u>
<u>Warmup Torque</u>	
OEM Torque	

Figure 2. Key Error Surfaces Effecting NO_x Deltas

Error surfaces which increased variation resulted in a wider distribution of errors, potentially both negative and positive. Error surfaces which contained a bias tended to offset all of the deltas either up or down. These results can aid in the interpretation of the results shown in Figure 1.

Method 1 shows a significant positive bias compared to the other two methods. The analysis results in Figure 2 indicate that Method 1 deltas are driven higher by a positive bias resulting from the Steady-State Exhaust Flow surface. The results given in the first paper in this series indicate that this is due to the fact that all of the exhaust flow measurement errors observed during model testing were positive. Methods 2 and 3 do not depend on accurate exhaust flow measurement, and are therefore not subject to this same positive offset.

Method 2 and 3 deltas, on the other hand are driven downwards by the Steady-State CO₂ surface. The results from the first paper indicate that CO₂ errors during model testing were always positive. As can be seen in Equations [2] and [3], positive CO₂ errors would result in a negative pressure on brake-specific values for these methods, while Method 1 would not be affected.

The torque and BSFC terms can also be seen as significant sources of variability in the final results. This highlights the importance of the accuracy of these ECM broadcast variables in the generation of accurate brake-specific emission values in the absence of an actual measured torque value.

NMHC EMISSIONS

The results of the Full Model run for BSNMHC are shown in Figure 3. These results are plotted against the "ideal" calculated NMHC value for the given reference NTE event. The NTE threshold is also shown for reference at 0.28 g/kW-hr. It should be noted that the PEMS could only measure THC, while the laboratory reference measurements used to generate the error surfaces were NMHC measurements. Therefore, a certain amount of positive error is anticipated for BSNMHC. It should be noted that for the PEMS, the THC measurement was multiplied by 0.98 to determine NMHC, which is the factor allowed in 40 CFR Part 1065 in the event the methane is not determined during measurement to allow calculation of NMHC.

As anticipated, all of the 95th percentile deltas were positive. It should be noted that the vast majority of the reference NTE event data was concentrated near zero NMHC emissions, and nearly all of the data was below 0.10 g/kW-hr. Similarly, much of the data generated during the model experiments was at levels very near zero NMHC. However, this result is not entirely unexpected considering that the model experiments were conducted using engines equipped with catalyzed DPFs, and that much of the reference NTE data is meant to represent similar DPF equipped vehicles.

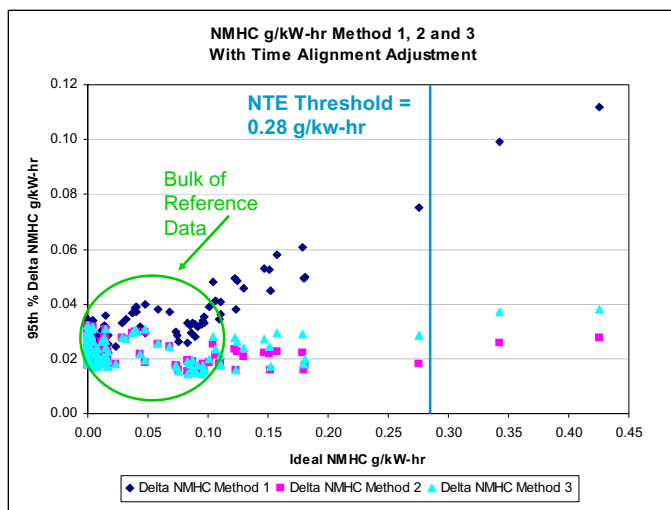


Figure 3. Full Model Results for NMHC

At these low levels, the deltas from all three calculation methods look similar to each other. At higher levels, near the NTE threshold for NMHC, Method 1 shows a high bias compared to the other two calculation methods. However, it should be noted that this represents only a small subset of the 195 reference NTE events. With the exception of 3 events, nearly all of the reference NTE data is far below the NTE threshold. This was also the case for the model data experiments.

A summary of the analysis of the key error surface drivers for NMHC deltas is given in Figure 4. The primary driver of NMHC deltas for all three calculation methods was the Ambient NMHC error surface. This surface resulted in nearly all of both the variation and bias effects for NMHC, and the bias was a strong positive offset for all three methods.

<i>Variation (Spread)</i>	<i>Bias (Offset)</i>
Method 1 Ambient NMHC Steady-State NMHC	Method 1 Ambient NMHC (+) Steady-State NMHC (+)
Method 2 Ambient NMHC Steady-State NMHC	Method 2 Ambient NMHC (+) Steady-State NMHC (+)
Method 3 Ambient NMHC Steady-State NMHC	Method 3 Ambient NMHC (+) Steady-State NMHC (+)

Figure 4. Key Error Surfaces for NMHC Deltas

This result is understandable given that the ambient HC environmental testing produced a significant high bias in the error surface. This bias was driven by the possibility of a HC FID auto-zero event happening in a region where ambient methane levels were higher than those observed during subsequent in-use measurements. Because the PEMS used ambient air as a zero gas, and because the ambient air filter used on the PEMS would

not block methane, such an occurrence would result in a positive bias in the results until the next auto-zero event. The magnitude of this error surface is larger than all other NMHC surfaces, and therefore it dominates the NMHC deltas. However, it should be noted that, as seen in Figure 3, this bias would not result in a threshold failure unless the test vehicle engine was already close to the NMHC threshold, which is typically not the case for DPF-equipped vehicles.

CO EMISSIONS

The results of the Full Model run for BSCO are shown in Figure 5. These results are plotted against the “ideal” calculated NMHC value for the given reference NTE event. For reference, the NTE threshold for CO is 26 g/kW-hr, which does not even appear on the chart at the axis scaling given to view the data.

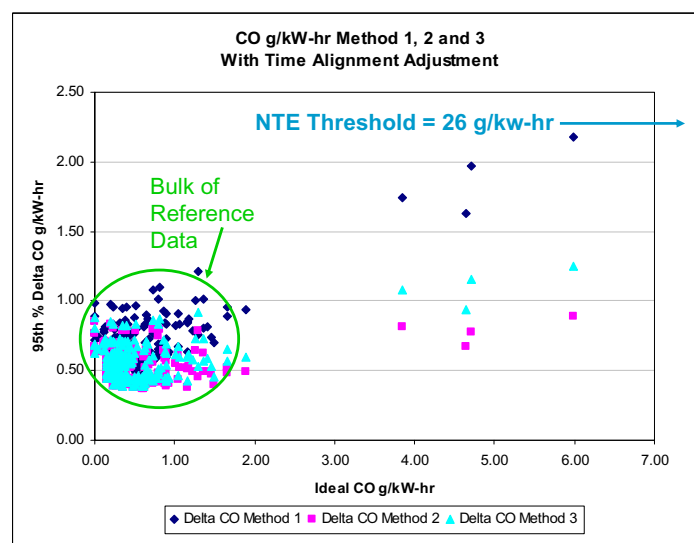


Figure 5. Full Model Results for CO

As can be seen in the data, there is a clear positive bias on all of the 95th percentile deltas for CO. This is the result of the fact that the PEMS uses a relatively low resolution NDIR detector for CO, which is in fact similar in resolution to the one used in the PEMS for CO₂. In all of the laboratory model experiments, a significant positive CO reading was observed even when the instrument was set to measure zero air through its sampling system. It should be noted however, that all of the reference NTE data was generally below 1 g/kW-hr, with the bulk of the data concentrated below 0.5 g/kW-hr. These levels are far below the NTE threshold of 26 g/kW-hr.

A summary of the analysis of the key error surfaces that drove CO deltas is given in Figure 6. The primary driver of CO errors was the Steady-State CO error surface, both in terms of variation and bias.

<i>Variation (Spread)</i>	<i>Bias (Offset)</i>
Method 1 Steady-State CO OEM Torque	Method 1 Steady-State CO (+)
Method 2 Steady-State CO OEM BSFC	Method 2 Steady-State CO (+) DOE BSFC (-)
Method 3 Steady-State CO CO_Time Alignment Warmup Torque	Method 3 Steady-State CO (+) CO_Time Alignment

Figure 6. Key Error Surfaces for CO Deltas

This result is again a reflection of the low resolution of the CO detector used in the PEMS. However, it should be noted that even with this offset of roughly 0.5 to 1 g/kw-hr, threshold failures are unlikely unless a vehicle was very near the CO threshold. Given the fact that the HDIUT vehicles will generally be diesel vehicles, and that most of these will be DPF-equipped, it is anticipated that test vehicles will generally be far below the CO threshold, such that this offset would not be a compliance issue.

MODEL VALIDATION

Although the results of the model run can be informative in explaining possible drivers of PEMS measurement error, the results cannot be utilized until the model is validated. Therefore, an extensive validation effort was carried out to insure that the model results could be considered representative of in-use testing.

The validation method, determined at the start of the program, involved performing a set of field experiments using the CE-CERT Mobile Emission Laboratory (MEL). The MEL was used to measure emissions from a test vehicle which was also equipped with a PEMS, so that parallel measurements could be made. This would allow generation of a set of measurement deltas which could be compared to the range of deltas generated by the model. If the bulk of these field deltas fell within the range of values defined by the model output, the model would be considered validated as representative.

It should be noted that the CE-CERT MEL does not incorporate any means of torque measurement, nor does it incorporate a means of fuel flow measurement, as both of these measurements would require very complicated equipment and time consuming custom adaptations to each vehicle. Therefore, this exercise could not validate the torque and BSFC terms within the model. As a result, the CE-CERT data could not be compared to the Full Model result. Therefore, a Validation Model result was generated with all of the torque and BSFC error terms turned off, in order to provide an appropriate data set for comparison to the CE-CERT field data.

CE-CERT – SWRI CORRELATION

A vital first step in this validation was to establish correlation between the CE-CERT MEL and the SwRI laboratory facilities. Because data from these two facilities was used to generate deltas for comparison to each other, it was essential that close correlation between the two facilities be demonstrated.

The CE-CERT MEL was driven to SwRI and positioned near the transient test cell used for generation of the model experiment data. This was done following the completion of data experiments on the first of the three program test engines, which was the DDC Series 60 14-liter engine. The exhaust system was configured to allow the exhaust to be easily switched between the SwRI CVS tunnel, and the CE-CERT MEL CVS tunnel. Three days of experiments were conducted, wherein each facility would make triplicate measurements of the exhaust for both transient and steady-state cycles on each test day.

The primary cycle used for correlation was a “transient NTE” cycle which was developed for this program. This cycle incorporated a set of 30 short NTE events, interspersed with short periods of light load operation outside the NTE zone. Cycle control and work measurement were performed by SwRI test cell systems in all cases. A summary of correlation results is given in Table 1.

Table 1. Transient NTE Cycle Correlation Data

Facility		Transient Emissions , g/kw-hr					
		THC	CH4	NMHC	CO	NO _x	CO ₂
SwRI	Mean	0.005	0.0001	0.005	0.075	2.683	725.8
	Stdev	0.001	0.0036	0.005	0.003	0.026	1.5
	Cvar-1SD			108.9%	3.8%	1.0%	0.2%
CE-CERT	Mean	0.001	0.002	-0.001	0.056	2.741	745.2
	Stdev	0.0004	0.0004	0.0005	0.001	0.019	3.9
	Cvar-1SD			59.0%	3.3%	0.7%	0.5%
Difference	%point	-65%	2336%	-117%	-24.2%	2.1%	2.7%
	%standard	-1.6%	1.3%	-2.9%	-0.1%	1.9%	n/a

In general, the correlation between the two facilities was within 2% for NO_x emissions. Both facilities demonstrated good test-to-test and day-to-day repeatability, with the exception of NMHC levels which were essentially at a near-zero level as a result of the catalyzed DPF used during the program.

It should be noted that this exercise represented a rare opportunity to observe two sets of CVS measurement equipment on an identical test article. The correlation was judged by the PEMS Steering Committee to be sufficient for validation to proceed.