CRC Report No. RW-107

Assessment of the Relative Accuracy of the PM Index and Related Methods

April 15, 2019

COORDINATING RESEARCH COUNCIL, INC.
5755 NORTH POINT PARKWAY • SUITE 265 • ALPHARETTA, GA 30022

The Coordinating Research Council, Inc. (CRC) is a non-profit corporation supported by the petroleum and automotive equipment industries. CRC operates through the committees made up of technical experts from industry and government who voluntarily participate. The four main areas of research within CRC are: air pollution (atmospheric and engineering studies); aviation fuels, lubricants, and equipment performance; heavy-duty vehicle fuels, lubricants, and equipment performance (e.g., diesel trucks); and light-duty vehicle fuels, lubricants, and equipment performance (e.g., passenger cars). CRC's function is to provide the mechanism for joint research conducted by the two industries that will help in determining the optimum combination of petroleum products and automotive equipment. CRC's work is limited to research that is mutually beneficial to the two industries involved. The final results of the research conducted by, or under the auspices of, CRC are available to the public.

CRC Project No. RW-107:

Assessment of the Relative Accuracy of the PM Index and Related Methods

Prepared for:

COORDINATING RESEARCH COUNCIL

April 15, 2019

Prepared by:

Robert Crawford
Rincon Ranch Consulting
2853 S. Quail Trail
Tucson, AZ 85730

and

Jim Lyons
Trinity Consultants, Inc.
3301 C Street, Suite 400
Sacramento, CA 95816

ACKNOWLEDGEMENTS

This work was funded by the Coordinating Research Council (CRC), Inc. Mr. Robert Crawford of Rincon Ranch Consulting, Tucson, AZ was the Principal Investigator (PI) responsible for the project management and for the statistical analysis presented here. Mr. Phil Heirigs, then of Trinity Consultants, Sacramento, CA, was the lead analyst in the determination of the PM indices. Mr. Jim Lyons of Trinity Consultants was the Co-Investigator. Mr. Jeff Jetter from Honda and Mr. Russ Lewis from Marathon Petroleum served as the CRC technical contacts for this project. Ms. Amber Leland represented CRC, the project sponsor.

CRC PROJECT NO. RW-107:

ASSESSMENT OF THE RELATIVE ACCURACY OF THE PM INDEX AND RELATED METHODS

Table of Contents

List of Figures

List of Tables

LIST OF ACRONYMS

AKI	Anti Knock Index
ASTM	American Society for Testing and Materials
CAFE	Corporate Average Fuel Economy
CO	Carbon Monoxide
CO_2	Carbon Dioxide
CRC	Coordinating Research Council
DBE	Double Bond Equivalent
DHA	Detailed Hydrocarbon Analysis
DOE	U.S. Department of Energy
EPA	U.S. Environmental Protection Agency
E0	Fuels that do not contain ethanol
E10	Fuels containing 10 vol% ethanol
E15	Fuels containing 15 vol% ethanol
E20	Fuels containing 20 vol% ethanol
E>0	Fuels containing ethanol at any level
EtOH	Ethanol Content
FID	Flame Ionization Detector
GC	Gas Chromatograph or Chromatography
GHG	Greenhouse Gas
MAE	Maximum Absolute Error
MS	Mass Spectrograph or Spectroscopy
MSD	Mass selective Device (in MS)
MW	Molecular Weight
NEDC	New European Driving Cycle
NOx	Oxides of Nitrogen
OESI	Oxygen Extended Sooting Index
PASCE	Particulate and Soot Correlation Equation
PEI	Particulate Evaluation Index
PM	Particulate Matter
PMI	Particulate Matter Index (ASTM DHA)
PMI-A	Particulate Matter Index (AVFL-29 DHA)
PN	Particle Number
PFI	Port Fuel Injection
R^2	Coefficient of Determination
RMS	Root Mean Square
RVP	Reid Vapor Pressure
SIDI	Spark-Ignited Direct-Injection
SP	Smoke Point
SS	Sum of Squares
SSI	Separation Systems, Inc.
SwRI	Southwest Research Institute
THC	Total Hydrocarbon
TSI	Threshold Sooting Index
US	United States
VP	Vapor Pressure

EXECUTIVE SUMMARY

Automotive engine technology has changed substantially in the past two to three decades under the influence of increasingly stringent emissions and fuel economy standards in the US and other countries. Recent actions in the US on Corporate Average Fuel Economy (CAFE) and Greenhouse Gas (GHG) emissions standards for model year 2017-2025 light-duty vehicles have stimulated the introduction of new engine technologies such as the Spark Ignition Direct Injection (SIDI) engine. Research has shown that Particulate Matter (PM) emissions from SIDI engines are generally higher than their Port Fuel Injected (PFI) predecessors, although some recent SIDI designs have demonstrated equivalent PM emissions to PFI. The composition and properties of gasoline fuels influence the formation and emission of particulates in both engine types.

Over the past 10 years, CRC, EPA, and other organizations have conducted research on the PM emissions of gasoline vehicles and the relationship to fuel properties. In such work, the Honda Particulate Matter Index (PMI) has found widespread acceptance as a measure of a gasoline's potential to contribute to PM emissions. The PMI is based on a profile of the individual hydrocarbons determined through a Detailed Hydrocarbon Analysis (DHA). CRC recently undertook the CRC AVFL-29 project to enhance the DHA procedure and permit it to identify more of the heavy compounds, which are of particular concern with respect to a gasoline's PM potential.

The ASTM D6729, D6730, and D6733 methods, as well as AVFL-29 DHA method, require specialized laboratory equipment and procedures for sample handling and analysis. While these methods can be employed in research settings, they may be beyond the capabilities of some laboratories and are likely to be too time consuming and costly for routine use in daily operations such as commercial fuel blending. As a result, a number of parties have proposed alternative PM indices relying on fuel properties that are more easily obtained and may be routinely available for other purposes. The objective of CRC Project RW-107 is to conduct a thorough evaluation of the performance of the different PM indices as predictors of vehicle emissions and a comparison of their performance to the Honda PMI using consistent sets of vehicle/fuel data.

Table ES-1 shows the eight PM indices evaluated in this work. Three indices – PMI, PMI-A and PEI – are based on the speciation of hydrocarbons in the fuel. The PMI-A index is defined by evaluating the Honda Equation for PMI using the DHA procedure developed under CRC AVFL-29. Fuel composition will be determinative of PM emissions and is directly accounted for in these indices. The remaining five indices rely on one or more bulk fuel properties to capture the correlation of the selected fuel properties to measured PM emissions.

Table ES-1
The PM Indices and Fuel Variables Required

PM Index	Fuel Variables
PMI-A	AVFL-29 DHA
PMI	ASTM D6729 DHA (or similar)
PEI	Aromatics speciation (C7 to C13)
Inverse SP (1/SP) OESI	Smoke Point Smoke Point Molecular Weight (MW) C, H, O wt%
PASCE	C/H Ratio E170°C
E130-E170 E150	E130°C, E170°C E150°C

To evaluate the performance of the eight indices in predicting PM emissions, three prior emissions studies were used as testbeds: the CRC E-94-2 and E-94-3 studies and the EPAct study conducted jointly by EPA, DOE, and CRC. The recent E-94-2 and E-94-3 studies tested a total of 12 and 4 SIDI vehicles using 8 and 12 gasoline fuels, respectively. The earlier EPAct study tested a total of 15 PFI vehicles using 27 gasoline fuels.

The PM indices were evaluated for the fuels used in the three emission studies and the performance of the indices in predicting PM emissions was compared. Three different metrics were used to evaluate predictive performance, including the overall goodness-of-fit and the sizes of the typical (root mean square) and the worst-case errors made in prediction.

Based on the consensus of the three performance metrics and the three emission studies, the following summarize the key findings and conclusions of the work:

1. The three indices based on detailed hydrocarbon speciation – PMI, PMI-A, and PEI – are the consistent high performers across all of the emissions studies. They benefit from detailed information on the hydrocarbon composition of fuels and can predict PM emissions well across a wide range of gasoline fuels.

2. CRC developed the AVFL-29 procedure for detailed hydrocarbon analysis with the goal of improving the speciation of the heaviest hydrocarbons, which are of particular concern for PM emissions. While the AVFL-29 procedure provides a

more complete assessment of hydrocarbon composition, that did not translate into improved predictive performance by PMI-A for the fuels and vehicles assessed here. PMI-A is as good a predictor of PM emissions as PMI, but it is not notably better. PMI-A should be a better predictor for fuels that contain atypically high or low amounts of the heaviest hydrocarbons or that explore the different ways that heavy hydrocarbons occur in gasoline.

3. In the E-94 studies, the correlation-based indices PASCE and E130-170 also perform well when the fuel groups are restricted to share the same level of oxygenation (i.e., fuels without ethanol or fuels containing 10 vol% ethanol). However, their performance is not as good for the EPAct fuels, where the complexity of the fuel matrix and the statistical requirements of the experiment likely make its fuels more dissimilar to the fuel datasets from which the correlation-based indices were developed.

4. All of the indices demonstrate a mathematical bias (a systematic error) in predicting PM emissions for sets of gasolines comprised of both E0 and E10-E20 fuels oxygenated with ethanol. The measured PM emissions of E10-E20 fuels are consistently higher and the emissions of E0 fuels are consistently lower than the values predicted by the overall correlation of the indices to emissions based on all fuels (E0-E20). This is true of both the speciation-based and the correlation-based indices. The average difference between E0 and E10 emissions (determined by their residuals from the correlation lines at equal PM index values) is small in a few cases, but more typically ranges from 10 to 37 percent depending on the index.

5. In their present form, five of the indices (PMI, PMI-A, PEI, PASCE, and E130-150) gave comparable accuracy for the E-94 fuels when the fuels share the same oxygenation level. This same performance may be achieved when the indices are applied to other fuels, provided that the fuels are generally similar to those considered here.

6. None of the PM indices is able to resolve the impact of ethanol on PM emissions for fuels in the E10-E20 range. Given this, the most important goal for future development work is to correct the bias so that at least some of the indices can be used for sets of gasoline fuels with varying ethanol content in the E10-E20 range.

1. INTRODUCTION

1.1 Background

Automotive engine technology has changed substantially in the past two to three decades under the influence of increasingly stringent tailpipe emissions and fuel economy standards in the US and other countries. Such pressures led early on to the replacement of carburetors with port fuel injection (PFI) systems to permit precise control of the air fuel mixture. More recently, actions in the US on Corporate Average Fuel Economy (CAFE) and Greenhouse Gas (GHG) emissions standards for model year 2017-2025 light-duty vehicles have stimulated the development of new engine technologies. Among them is the Spark Ignition Direct Injection (SIDI) engine that uses sophisticated computer control to inject atomized gasoline into the cylinder in the effort to maximize engine output and minimize fuel use.

Gasoline engines will generate fine particulate matter (PM) in the exhaust in circumstances where liquid fuel droplets are combusted or where air-fuel mixing is incomplete. In conventional engines (e.g., PFI fuel systems), this occurs mainly during cold-start operation and will greatly diminish once the engine becomes fully warm. There are specialized circumstances – such as the operation of motor vehicles in extremely cold climates – where PM emissions from gasoline vehicles can be a significant source of fine particulates in urban areas, but such circumstances are uncommon.

With the introduction of SIDI engine technology, atomized gasoline droplets are now combusted in a wider range of operating modes. Cold-start operation continues to be when the largest amount of PM is emitted, but PM emissions can also occur throughout the drive cycle. Recent research has shown that PM emissions from SIDI vehicles are generally higher than from PFI vehicles (although some recent SIDI designs have demonstrated comparable emissions). In addition to engine and emission control design, the composition and properties of gasoline fuels have been shown to influence particulate emissions in both engine types.

1.2 Objectives

Over the past 10 years, CRC, EPA, and other organizations have conducted research on the PM emissions of gasoline vehicles and the relationship to fuel properties. In this work, the Honda Particulate Matter Index (PMI) has found widespread acceptance as a measure of a gasoline's potential to contribute to PM emissions. The PMI is based on a profile of the individual hydrocarbons in a fuel as determined through a Detailed Hydrocarbon Analysis (DHA), a typical method for which is the ASTM D6730 procedure as used

originally in its development. Variants of the ASTM DHA procedure also have been used with the PMI.

The DHA is a sophisticated laboratory procedure that uses a Gas Chromatograph (GC) and Flame Ionization Detector (FID) to determine the presence and quantity of more than 200 individual hydrocarbons in a fuel. Because the heaviest hydrocarbon compounds may be disproportionately responsible for PM emissions, CRC recently undertook the CRC AVFL-29 project to enhance the DHA procedure and enable the identification and quantification of more of the heavy compounds in a fuel.

The ASTM and AVFL-29 DHA methods require specialized laboratory equipment and procedures for sample handling and analysis. While these methods can be employed in research settings to support use of the Honda PMI, the methods may be beyond the capabilities of some laboratories and are likely to be too time consuming and costly for routine use in daily operations such as commercial fuel blending. As a result, a number of parties have proposed alternative PM indices based on fuel properties that are more easily obtained and may be routinely available for other purposes. These alternative indices range from one that relies on a simplified profile of gasoline hydrocarbons to several that are based on commonly available bulk properties and/or flame behavior measured in laboratories. The indices have been shown by their authors to correlate well with the Honda PMI and to adequately predict PM emissions based on the generally small datasets of fuels and engines used in their development.

Because the alternative indices have often been developed using different vehicles and fuels, CRC concluded that a thorough evaluation of the performance of the indices and a comparison of their performance to the Honda PMI was needed using consistent sets of vehicle/fuel data. This evaluation was the objective of CRC Project RW-107, which is reported here.

In this work, eight different PM indices are applied to the experimental fuels tested in three different emission studies: the CRC E-94-2[1] and E-94-3[2] projects and the EPAct[3] study. These fuel sets are different than the ones from which the indices were originally developed. The CRC studies involved a test fleet of 12 and 4 SIDI vehicles with a total of 8 match-blended and 4 splash-blended fuels including E0 and E10 gasoline. The EPAct study involved a test fleet of 15 PFI vehicles with a total of 27 experimental fuels, ranging from E0 to E20 gasoline. The PM indices are evaluated for the experimental fuels in the studies and then applied in the prediction of PM emissions from the vehicles over Phase 1 of the LA92 test cycle. Three different metrics are used to measure the performance of the indices in predicting emissions. The result of the comparative analysis is a ranking of the indices in order of performance.

[1] https://crcao.org/reports/recentstudies2017/E-94-2/CRC_2017-3-21_03-20955_E94-2FinalReport%20Rev1b.pdf
[2] https://crcao.org/reports/recentstudies2018/E-94-3/CRC-E-94-3_Final%20Report_2018-06-26.pdf
[3] https://www.epa.gov/moves/epactv2e-89-tier-2-gasoline-fuel-effects-study

1.3 Report Organization

The report is organized as follows. Section 2 describes the three emission studies that are used in this work as the basis for the comparative evaluation of the PM indices. Section 3 reviews the PM indices examined in the study, including their mathematical form and the key characteristics of the datasets from which they were developed. Section 4 presents a summary of the analytical methodology used to determine the PM index values for the experimental fuels and to evaluate index performance in predicting emissions. Section 5 examines the performance of the PM indices for E0 gasoline (i.e., not containing oxygenates). Section 6 examines index performance on gasolines that range from E0 to E10 in ethanol content. Section 7 examines index performance for the full range of E0 to E20 gasolines considered in the EPAct study. Section 8 presents the conclusions of the study and recommendations for further work.

Section 9 presents a list of technical references. Within the report, citations to the technical literature are given in the format (Author Date); full citations will be found in Section 9. The references include the documents used in this report as the definitive source on each PM index.

Four appendices have been prepared to document the analysis performed in this work. The first three appendices are intended as a reference to the detailed results and are presented without commentary. Appendix A compiles the PM index performance results for the group of E0 gasolines in the form of tables and figures. Appendix B provides a similar compilation of performance results for the group of E0-E10 gasolines. Appendix C provides the same for the group of E0-E20 gasolines that were considered in the EPAct study. Appendix D describes the methodology used to evaluate the statistical uncertainties in the performance metrics used here.

2. PARTICULATE EMISSIONS DATASETS AND EXPERIMENTAL FUELS

Three recent emissions studies are used in this work as platforms to test the performance of the PM indices for predicting PM emissions. The studies are:

- The CRC E-94-2 and E-94-3 projects, which were conducted between 2016 and 2018 to determine how gasoline characteristics influenced the PM emissions of SIDI engines; and

- The EPAct study, completed in 2013 as a joint research project of EPA, DOE and CRC, which examined how selected bulk properties of gasoline influenced the PM emissions of PFI engines.

The following sections highlight the major characteristics of the vehicles and fuels tested in each study.

2.1 CRC E-94-2 Study

The CRC E-94-2 study investigated the particulate and gaseous emissions of a test fleet of 12 vehicles equipped with SIDI engines (see Table 2-1). Ten of the vehicles were equipped with 4-cylinder engines, split equally between turbocharged and naturally aspirated air induction systems, while two were equipped with 6-cylinder engines. As a group, the vehicles were considered to be representative of the range of SIDI models available in the US.

Eight different gasolines were blended for the study according to a design matrix based on octane rating (AKI), ethanol content (EtOH), and PMI level. Two different AKI levels were targeted: 87 AKI, representative of regular grade gasoline; and 93 AKI, representative of premium grade in much of the US. Four fuels were blended within each octane level to fill a 2 x 2 matrix of ethanol content (E0 and E10) and PMI level (Low PMI $\leq$ 1.40 and High PMI $\geq$ 2.4). The characteristics of the fuels, as blended, are given in Table 2-2.

To guide the blending of each fuel, CRC provided specifications and tolerances for a number of fuel properties[4]. For example, total aromatics content was targeted at 25 $\pm$ 2 vol% and further specifications were given for benzene and olefins content. Low and High PMI fuels were created by varying the composition of aromatic compounds between light

[4] See Section 3.1 Fuels, and Tables 1 and 2, of the CRC E-94-2 Report (2017) for further information.

Table 2-1
Test Fleet used in the E-94-2 Study

Vehicle	Engine Type
2011 Chevrolet Equinox	2.4L Naturally Aspirated, I4
2013 Chevrolet Malibu	2.0L Turbocharged, I4
2013 Chevrolet Malibu	2.5L Naturally Aspirated, I4
2013 Ford F150XL	3.5L Turbocharged, V6
2013 Honda Accord	2.4L Naturally Aspirated, I4
2013 Hyundai Santa Fe	2.0L Turbocharged, I4
2013 Hyundai Santa Fe	2.4L Naturally Aspirated, I4
2015 Lexus NX200t	2.0L Turbocharged, I4
2014 Mazda Mazda6	2.5L Naturally Aspirated, I4
2013 Mercedes-Benz GLK350	3.5L Naturally Aspirated, V6
2010 Nissan Juke	1.6L Turbocharged, I4
2012 VW Jetta GLI	2.0L Turbocharged, I4
Source: Table ES-2. CRC E-94-2 Report (2017).	

Table 2-2
Test Fuels used in the E-94-2 Study

Fuel Letter	AKI	EtOH[a/] (vol%)	PMI
A	87.2	9.55	1.42
B	87.1	9.56	2.65
C	87.9	0	1.40
D	88.2	0	2.61
E	93.6	9.56	1.28
F	93.7	9.51	2.54
G	93.8	0	1.26
H	94.1	0	2.49[b/]

[a/] Based on ASTM D4815.

[b/] The PMI value for Fuel H has been corrected from that published to reflect the re-analysis of the retained sample at the start of the E-94-3 program. The revised value, determined by Lab C, is consistent with the PMI value also determined by Lab C for the splash-blended Fuel H-C10. The PMI values for the E-94-2 fuels published in Appendix A of the CRC E-94-2 Report (2017) are based on averages for three laboratories.

Source: Table 3. CRC E-94-2 Report (2017).

(< C10) and heavy (C10+) molecules. Additional, but light-handed constraints were imposed on properties such as RVP and the T50, T90, and FBP boiling points. Fuel analyses were performed at four different laboratories to assure that the fuels met the CRC target specifications.

The overall objective in blending was to create fuels that, to the extent possible, would be representative of commercial gasoline across the typical range of octane levels, ethanol contents and PMI values. The properties of the blended fuels were checked to assure that each met the specifications for commercial gasoline. This was considered important to demonstrate that the blended fuels are, in fact, commercial gasolines and could be sold in the US market. For this reason, the set of match-blended E-94-2 fuels are believed to be the most representative of US gasolines among the fuel sets considered in the three emissions studies. However, they are only eight in number and do not explore the full range in variation for fuel characteristics that influence PM emissions

The E-94-2 fuels are termed "match blended" as the fuel blender was instructed to match the specified targets within the given tolerances; the label E10-M notes their match-blending and is used to identify these fuels in figures. Specifically, the E10 fuels for each AKI and PMI level were not blended from the E0 fuel of the same level, but from a revised base fuel formulation so that the E10 fuel would meet the design targets once ethanol was added at the 10 vol% level. This process matches how most gasoline is blended in the US today. Thus, the gasoline hydrocarbons in the E10-M fuels differ somewhat from the corresponding E0 fuels and are, in general, of lower octane, lower volatility, and higher PMI before the blending with ethanol.

Particulate and gaseous emissions of the vehicles were measured at the Southwest Research Institute (SwRI) laboratory in San Antonio, TX, using the LA92 test cycle. Emissions were reported for the individual phases of the cycle and as LA92 weighted-averages. Both particulate mass (PM) and particle number (PN) were measured along with subcomponents (Elemental and Organic Carbons). Gaseous emissions were measured as well, including THC, CO, NOx, and CO_2.

2.2 CRC E-94-3 Study

The CRC E-94-3 study used a subset of four vehicles that participated in the E-94-2 study to examine a specific question regarding the emissions of E10 fuels. The results of the E-94-2 study demonstrated that PM emissions of the E10 fuels were higher than would be expected from their PMI values and higher than observed from E0 fuels of comparable PMI. Because the gasoline hydrocarbons in the match-blended E10 fuels differed from those in the E0 fuels, the E-94-3 study was conducted to determine whether the method of blending was responsible for the higher emissions.

Four additional E10 fuels were created by splash-blending 10 vol% ethanol into each of the four E0 fuels tested in E-94-2 (see Table 2-3). The label E10-S notes their splash-blending and is used to identify these fuels in figures. In this case, the gasoline hydrocarbons of the splash-blended E10 fuels were identical to those in the E0 fuels except

for dilution. No effort was made to control other properties of the resulting E10 fuels, which were generally more volatile and of higher octane than their E10 counterparts in E-94-2.

Emissions testing was performed at the SwRI laboratory using the LA92 cycle. Both PM and PN emissions were measured, along with a range of gaseous pollutants. Four vehicles from the E-94-2 program were tested[5] using the four splash-blended E10 fuels created for the study. The vehicles were not retested using the E0 and match-blended E10 fuels of the E-94-2 study. Instead, the prior testing of the vehicles was merged with the new testing to create a data representing the PM emissions of 4 vehicles on 12 fuels.

Table 2-3
Test Fuels used in the E-94-3 Study

Fuel	Base Fuel	AKI	EtOH[a/] (vol%)	PMI
C-E10	C	91.5	9.44	1.28
D-E10	D	91.1	9.71	2.45
G-E10	G	96.4	9.75	1.17
H-E10	H	96.0	9.88	2.32

[a/] Based on ASTM D4815.
Source: Table ES-1. CRC E-94-3 Report (2018).

The later sections of this report examine the performance of the PM indices for a sequence of fuel groups beginning with E0 fuels in each study, then the E0 to E10 fuels and, finally, all fuels from E0 up to E20. For the E0 fuel group, the results shown for the E-94-3 study are simply a subset of the results for the E-94-2 study. The four E0 fuels are the same and the same four E-94-3 vehicles are contained in the larger test fleet of 12 vehicles in the E-94-2 study. Thus, differences in the PM index performance results between the two E-94 studies for E0 gasoline are reflective only of the differing emissions performance of the 12 and 4 vehicle test fleets.

For the group of E0-E10 fuels, the results for the E-94-2 and E-94-3 studies are based on different fuels. The E-94-2 results are based on the E0 and match-blended E10 fuels (8 in total), while the E-94-3 results are based on the E0, match-blended E10, and splash-blended E10 fuels (12 in total) for 4 of the 12 vehicles. Differences in the PM index performance results between the two studies reflect a combination of differing index performance for splash- versus match-blended E10 fuels and the differing emissions performance of the 12 and 4 vehicle test fleets.

[5] 2011 Nissan Juke, 2013 Chevrolet Malibu (naturally aspirated), 2015 Lexus NX200t, and 2013 Mercedes-Benz GLK350.

2.3 EPAct Study

The EPAct study, a joint research project involving EPA, DOE, and CRC, was conducted in phases between 2007 and 2013 to understand how changes in vehicle technology since the 1990s had changed the relationship between vehicle emissions and gasoline properties. This study uses the EPAct Phase 3 emissions results, which are documented in three reports: EPAct (2013a) Analysis of Data, EPAct (2013b) Program Design and Data Collection, and EPAct (2010) Fuel Blending.

The study focused on five key characteristics of gasoline—aromatics and ethanol content, RVP, and the T50 and T90 boiling points—and used a fleet of 15 test vehicles from model year 2008. The vehicles included passenger cars and light trucks equipped with a variety of engines types and sizes[6], but all of the vehicles had PFI fuel systems. The PFI vehicles produced much lower levels of PM emissions on their gasoline fuels than the SIDI vehicles of the two E-94 studies. The lower level of PM emissions is a key factor distinguishing EPAct and the two E-94 studies.

The EPAct fuels were blended to meet a design matrix of five variables as shown in Table 2-4. Aromatics content was considered at two levels: 15 vol% (a low aromatics fuel) and 35 vol% (a high aromatics fuel). Ethanol was considered at four levels: at the E0 and E10 levels, which are commonly found in the US market, and at the higher E15 and E20 levels. Two RVP levels were considered to include both summer and winter gasolines. Five and three levels were considered for T50 and T90, respectively.

Table 2-4
Design Levels for Fuel Properties of EPAct Fuels

Factor	Levels	Design Values		
		Low	Middle	High
Aromatics (vol%)	2	15		35
Ethanol (vol%)	4	0	10, 15	20
RVP (psi)	2	7		10
T50 (°F)	5	150	165, 190, 220	240
T90 (°F)	3	300	325	340
Source: Table 2. EPAct (2013a).				

In total, 27 experimental fuels were created and tested for emissions in the fleet of 15 vehicles (see Table 2-5). The 27 fuels do not represent all combinations of the design levels because an attempt to do so would involve too many fuels (240) and many of the combinations do not exist among commercial fuels or are infeasible to blend. The final

[6] See Table 1 of EPAct (2013a).

selection of 27 fuels was the result of a process that balanced research goals and the feasibility of fuel blending with the experiment design.

Table 2-5
Design Characteristics of the 27 EPAct Fuels

Fuel	Aromatics (vol%)	Ethanol (vol%)	RVP (psi)	T50 (°F)	T90 (°F)
1	15	10	10	150	300
2	15	0	10	240	340
3	15	10	7	220	300
4	15	10	10	220	340
5	35	0	7	240	300
6	15	10	7	190	340
7	15	0	7	190	300
8	15	0	10	220	300
9	35	0	10	190	340
10	35	10	7	220	340
11	35	10	10	190	300
12	35	10	10	150	340
13	35	0	7	220	340
14	15	0	7	190	340
15	35	0	10	190	300
16	35	10	7	220	300
20	15	20	7	165	300
21	35	20	7	165	300
22	15	20	10	165	300
23	15	20	7	165	340
24	15	20	10	165	340
25	35	20	10	165	340
26	35	15	10	165	340
27	15	15	7	220	340
28	35	15	7	220	300
30	35	10	10	150	325
31	35	20	7	165	325

Source: Table 3. EPAct (2013a).

The EPAct study was specifically designed to understand the effect of fuel property changes on vehicle emissions. The study authors noted that: "[a] critical point about the design of the program is that the properties of the test fuel are assigned so as to span the boundaries of in-use fuel properties. This approach is designed specifically to provide a basis for the development of statistical models capable of predicting emissions for the majority of in-use fuels. …Test fuel parameter ranges were originally drafted to span roughly the 5th to 95th percentiles of survey results for US gasolines."[7]

Given this approach, the values for the five design variables are, individually, representative of the maximum range (boundaries) of fuel property values found in US gasolines. EPAct also includes fuels having atypical combinations of bulk properties—such as low T50 combined with high T90 (and vice versa)—in order to meet the requirements of the experiment. The blending used hydrocarbon streams commonly found in refineries, sometimes fractionated (as in the E-94 studies and most studies involving specialty fuels), but never individual hydrocarbon compounds in meeting the fuel property targets. Because they were blended to explore the maximum range of bulk properties and included atypical combinations of some properties, the EPAct fuels are likely to be more dissimilar to the fuel datasets from which the correlation-based indices were developed than the fuels in the two E-94 studies. The EPAct fuels may also be less representative of *typical* commercial gasolines found in the market.

Five of the PM indices considered here use individual fuel properties (e.g., the C/H ratio and selected distillation points) to predict PM emissions. Such indices rely on correlations between bulk fuel properties and the hydrocarbon compositions of fuels for their predictive power. Indices based on hydrocarbon speciation benefit from direct information on fuel composition and do not rely on property correlations. In the EPAct fuels, it is likely that the correlations between bulk properties and composition have been perturbed, to at least some extent, from the correlations typical of commercial gasoline and likely also found in the original datasets from which the correlations were developed. As shown in later sections, the correlation-based indices do not perform as well as the speciation-based indices in predicting PM emissions for the EPAct study.

2.4 Availability of Retained Samples

Three of the PM indices considered in this study require fuel property information that was not determined for the three emissions studies. The additional information includes the DHA performed according to the AVFL-29 procedure and the measured smoke points of the fuels. This information was obtained for the two E-94 studies through new analysis of the retained samples for their 12 fuels. The EPAct study was completed more than 5 years ago and retained samples of its fuels no longer exist. Thus, 3 of the indices[8] (PMI-A, 1/SP, and OESI*-DHA MW) cannot be evaluated for the EPAct study.

[7] EPAct (2013a), pp. 14-15.
[8] See Section 3 and Table 3-1, specifically, for definition of the indices.

3. THE PM INDICES

This section describes the PM indices that are examined in this study. The discussion defines the indices and summarizes how they were developed, including the chief characteristics of the fuels and vehicles or engines used in emissions testing. References given in each subsection identify the document used as the definitive source on the PM index.

The first two indices—PMI and PMI-A—are distinctive in that they define the potential of a fuel to form PM without the use of empirically determined constants. Instead, based on empirical findings from prior research, they posit that the contribution of a hydrocarbon species to PM emissions increases with its double-bond equivalent (DBE) value and decreases with its vapor pressure (VP) in proportion to the ratio (1+DBE)/VP. PMI and PMI-A require detailed speciation of the hydrocarbons in the fuel, using ASTM D7629 or similar for PMI and the AVFL-29 DHA procedure for PMI-A, to determine the weights by specie.

The third index, PEI, is an empirical index that relies on a speciation of aromatic hydrocarbons by carbon number, which can be derived from DHAs or by methods such as PIONA (ASTM D8071). It contains empirical constants to define the contributions of the aromatic carbon number groups to PM emissions potential, but it does not rely on correlations between fuel properties and composition. Along with PMI and PMI-A, the speciation-based indices pose a need for sophisticated laboratory procedures to speciate the hydrocarbon compounds in a fuel.

The remaining five indices are different in that they attempt to provide a simplified method for estimating a fuel's potential to form PM that does not require hydrocarbon speciation. These correlation-based methods generally rely upon one or more predictor variables for relevant characteristics of a fuel that should be correlated to the fuel's hydrocarbon composition and potential to form PM. A common practice is to estimate an empirical equation that predicts the PMI values of selected test fuels. The ability of the resulting index to predict PM emissions is then validated using emissions test data on one or more vehicles or engines.

Table 3-1 compiles a list of the PM indices examined in this study along with the fuel variables involved and the laboratory procedures used for their determination. The indices are listed in descending order of the sophistication of the laboratory procedures required for their application, beginning with the three speciation-based indices. For the correlation-based indices, the range of predictor variables and laboratory procedures varies from simple (one or more points on the distillation curve) to more complex (the determination of smoke point and molecular weight).

Table 3-1
The PM Indices: Fuel Variables and Analytical Procedures

PM Index	Fuel Variables	Analytical Procedures
PMI-A	AVFL-29 DHA	AVFL-29 Procedure
PMI	ASTM DHA	ASTM D6729
PEI	Aromatics Speciation (by carbon number)	PIONA or ASTM DHA
Inverse SP (1/SP) OESI*	Smoke Point Smoke Point Molecular Weight (MW) C, H, O Wt%	ASTM D1322 ASTM D1322 ASTM DHA ASTM ASTM D5291, D5599
PASCE	C/H Ratio E170°C	ASTM D5291 ISO 3405 / ASTM D86
E130-E170 E150	E130°C, E170°C E150°C	ISO 3405 / ASTM D86
Note: The asterisk in OESI* indicates that it is calculated from unscaled smoke point values, rather than scaled values that would permit comparisons across laboratories.		

3.1 PMI

The PMI index, defined in Aikawa 2010, is an index for the propensity of fuels to contribute to PM emissions. It is based on the ASTM DHA procedure and the Honda PMI Equation to estimate the PM potential of each hydrocarbon in the fuel from their DBE and VP values. In this report, the index is termed "PMI" to distinguish it from "PMI-A", which is calculated in the same way but with inputs from the AVFL-29 DHA procedure.

Aikawa 2010 found that hydrocarbon compounds with high boiling points and correspondingly low VPs and high DBEs were associated with increased PM emissions. It posited that the PM potential for hydrocarbon compound "i" would be proportional to the ratio $(1 + DBE_i) / VP_i$. The weighted summation of this ratio across the hydrocarbons present in a fuel defined an index number that was shown empirically to be proportional to the PM emissions generated by combustion of the fuel. The Honda Equation for PMI is given in Eq. 3-1 below, where the summation is over the individual hydrocarbons "i" in the fuel. The ratio inside the summation is often called the "i-term" and is taken to be the measure of each compound's relative contribution to the total PM emissions of the fuel.

$$\text{PMI} = \sum_i Wt_i * \left(\frac{1+DBE}{VP \text{ at } 443K} \right) \qquad \text{(Eq. 3-1)}$$

The DBE value is a characteristic determined by the chemical structure of each compound. The vapor pressure is that measured at 443K (170° C). The Appendix to Aikawa 2010 gives the definition of DBE and an empirical formula that can be used to estimate VP_{443k} from the boiling point. When the index is evaluated for fuels containing ethanol, an empirical value for VP_{443k} must be used in the calculation as its vapor pressure is very high in comparison to gasoline hydrocarbons.

It is important to recognize that this index contains no empirical constants. Aikawa 2010 says that the unity ("1") term in the numerator was a free choice in defining the index and was chosen as unity to set the emissions generated by saturated compounds (DBE=0) as the reference level for the index. But none of the terms in Eq. 3-1 is determined empirically, as is typically the case in other indices considered here.

Aikawa 2010 validated the usefulness of the PMI by measuring particle number (PN) emissions from one test vehicle (a 2.3L port fuel-injected turbocharged engine) over the New European Driving Cycle (NEDC). Nine test fuels were employed, including the US certification fuel (indolene) and eight fuels created by adding ethanol or one of 7 different hydrocarbons. The ethanol fuel was blended at the E21 level, while the eight other fuels contained the added hydrocarbon at nominally 10 % wt. Thus, the test fuels are predominantly E0 with one E21 fuel above the level found commercially in the US. Further verification was provided using a range of different engines, driving cycles, and fuels including commercial gasolines. In all cases, the observed particulate emissions were found to vary with the PMI value over a wide numerical range (1.0 to 4.0) for the index. This range extends well beyond the levels thought to be representative of commercial gasoline in the US

Since its publication, the PMI has become widely accepted in fuels research as a primary measure for the PM emissions potential of automotive gasoline. Here, it is used as a benchmark or a "gold standard" against which the performance of the other PM indices can be compared. Such use does not imply that the index is a perfect measure of PM formation potential but, rather, that it is the best available index for that purpose at present.

The PMI values used here are those published for the fuels in the E-94-2, E-94-3 and EPAct studies in the corresponding reports. The PMI values were calculated by SwRI from ASTM DHAs performed at the SwRI laboratory in San Antonio, TX. The DHA method employed at SwRI has evolved over time. The DHAs performed for the EPAct study (published in 2013) quantified hydrocarbons through the C12 group. For the two E-94-2 and E-94-3 studies (published in 2017 and 2018, respectively), the DHAs quantified hydrocarbons through the C15 group, albeit with declining rates of compound identification at high carbon number.

3.2 PMI-A

The PMI-A index is defined here as an alternative to the PMI in which a modified DHA procedure is used to extend the speciation of gasoline hydrocarbons to higher carbon

numbers with a higher rate of successful identification of heavy hydrocarbon compounds. The development work was performed by Separation Systems Inc. (SSI) in Gulf Breeze, FL. That work and the enhanced procedure are described in the CRC AVFL-29 Report (2018).

A number of ASTM methods are in use for DHA including ASTM D6729, ASTM D6730, and ASTM D6733. These methods differ in terms of the technical details associated with the analysis, but not the fundamental analytical approach, which is based on gas chromatography (GC). The GC method utilizes a capillary column to separate the individual chemical compounds in a sample so that they can be identified and then quantified by a flame ionization detector (FID) placed at the end of the column. The identification of a specific compound is made by comparing the amount of time required for each to elute from the column and reach the detector. The amount of the compound present in the sampler is determined by the output signal from the FID. The time required to pass through the column is referred to as the retention time.

As stated in the ASTM specifications that apply to DHA analysis:

> *Each eluting component is identified by comparing its retention time to that established by analyzing reference standards or samples under identical conditions. The concentration of each component in mass percent is determined by normalization of the peak areas after correction of selected components with detector response factors. The unknown components are reported individually and as a summary total. (ASTM D6729)*

One concern with use of DHA in PMI calculations is that the higher molecular weight compounds that may be of particular interest in terms of PM formation tend to elute late in the GC process. Their elution may overlap with the elution of other compounds making their identification and quantification impossible based on retention time.

In light of this, CRC conducted project AVFL-29 with the goal of reducing the fraction of a gasoline sample that remains unidentified by individual compound. The AVFL-29 approach was developed using a mass spectrometer (MS) to improve the ability of the GC based approach to identify specific higher molecular weight compounds based on their retention time. MS identifies specific molecules based on the fragments produced after their ionization when they pass through a magnetic field under specific conditions. The fragmentation patterns of specific compounds after ionization are generally known and can be used in to identify specific chemical compounds. After identification of the compounds and their retention times using MS, that information was used to "enhance" the ability of the GC procedure to quantify higher molecular weight compounds. The enhanced procedure is estimated to reduce the fraction of a gasoline sample that remains unidentified from approximately 5% with the ASTM DHA to 0.5% with AVFL-29[9]. This is of particular importance with respect to PM emissions, as the heavy hydrocarbon compounds

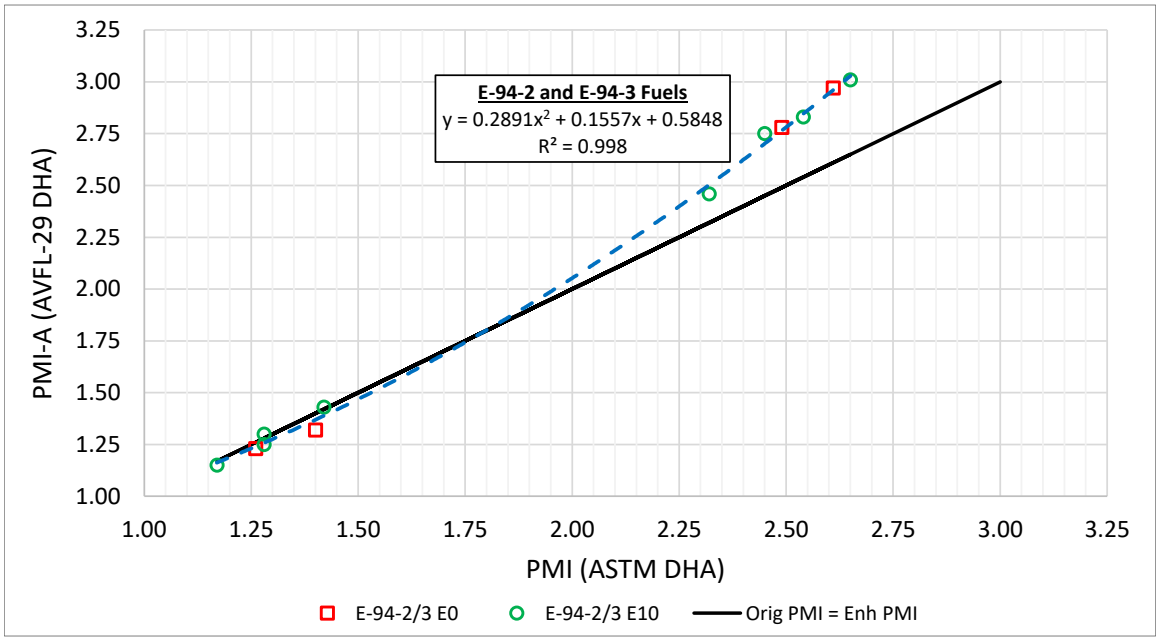
[9] Relative retention times are listed for all components identified by AVFL-29. This allows a laboratory to perform the enhanced DHA with only an FID detector.

not identified by the ASTM DHA can make disproportionate contributions to PM emissions even if they make up only a small portion of a fuel.

For this study, SSI applied the AVFL-29 procedure to retained samples of the twelve E0 and E10 fuels from the E-94 studies to determine PMI-A their values based on the Honda Equation (Eq. 3-1 above). Analysis was conducted to understand the performance of the procedure for these fuels and to resolve questions regarding the comparison of PMI-A values to the PMIs determined by SwRI. Retained samples of the earlier EPAct fuels no longer exist, so the PMI-A could not be evaluated for them.

Figure 3-1 shows the correlation of the SwRI PMI to the SSI PMI-A with the quadratic correlation equation given by the dotted blue line[10]. The AVFL-29 DHA procedure has relatively little effect on the PMI values of the low PM fuels (at the left with index values below 1.50) because these fuels contain few heavy hydrocarbons. But the procedure has a large effect on the PMI values of the high PM fuels (at the right with index values above 2.25). There, the calculated PMIs are increased compared to the PMI values and lie well above the line of equality (solid black) between the two indices. All data points lie close to the correlation line with the Coefficient of Determination (R^2) equal to 0.998.

Figure 3-1
Correlation of PMI-A (AVFL-29 DHA) to PMI (ASTM DHA)



The very high R^2 value indicates that the two indices are equivalent measures for this group of fuels. Equivalence means that one set of index values could be transformed into the other set of values with good accuracy and little loss of predictive power. In fact, the later analysis will indicate that the Enhanced and PMIs perform equally well for the E-94-2 and

[10] A curvilinear correlation is shown here to emphasize the differential effect on high PM fuels. A linear correlation line would give a similar fit to the individual data points with a comparably high R^2 of 0.997.

-3 fuels. This result says that the prevalence of the heaviest hydrocarbons in these fuels, which are detected only in the AVFL-29 DHA, is highly correlated with the prevalence of other heavy hydrocarbons that are identified in the ASTM DHA procedure. In this case, PMI can have comparable accuracy in predicting emissions even though it does not directly contain information on the heaviest hydrocarbons that PMI-A contains

This outcome should not be a surprise, as the E-94-2 fuels were blended from selected refinery streams with attention given to matching the typical characteristics of commercial gasoline. Both Low and High PMI values were targeted, but the experiment did not attempt to explore different ways of achieving High PMI values. The E-94-3 fuels were then splash-blended from the E0 fuels of the E-94-2 study. If only one heavy hydrocarbon stream were used to create the High PMI fuels, then there will be a fixed ratio between the heaviest hydrocarbons identified only in the AVFL-29 procedure and the heavy hydrocarbons identified in the ASTM procedure. The PMI's information on the heavy hydrocarbon content would be exactly correlated to the (unknown) content of the heaviest hydrocarbons identified only in the AVFL-29 procedure. The same would be nearly true if two different, but similar heavy hydrocarbon streams were used in the blending. For such fuels, and other gasolines like them, the two indices are likely to perform equally well in predicting emissions.

For other fuels—particularly fuels that are atypically high or low in the heaviest hydrocarbons in comparison to other heavy compounds—the PMI-A may well be a better predictor of PM emissions. This would be most easily seen in a dataset of fuels that had purposely been blended to explore the different ways in which heavy hydrocarbons occur in gasoline. In this case, a number of different heavy hydrocarbon streams would be used to create a variety of High PMI fuels, which would likely have varying ratios between the heaviest hydrocarbons detected only by the AVFL-29 procedure and the heavy hydrocarbons detected by the ASTM procedure. For such a dataset, the PMI-A should prove to be a superior predictor, since its DHA carries additional information on heavy hydrocarbons that is not available to the PMI.

A number of comparisons were made during the determination of the PMI-A values to assure the quality of the result. One was of the ethanol content, among the most volatile components of the fuels, as determined by the AVFL-29 DHA (SSI) and the ASTM DHAs (SwRI). As Table 3-2 shows, the E10 fuels used in the two E-94 studies demonstrated nearly the same ethanol contents when the retained samples were analyzed by SSI compared to when the fuels were originally analyzed by SwRI. The percentage changes ranged from -4% to +1% on a relative basis, which is generally too small to be reliably measured by the procedures. Also, there are many differences between the DHA procedures that can contribute to differences, including detailed differences in equipment between the labs and procedural differences in the integration of the ethanol peaks (such as manual integration versus auto-integration).

It remains the case that PMI and PMI-A were determined by different laboratories and at different times (2-3 years apart) and could be affected by the existence of systematic

differences other than simply the use of different DHAs. The precision of the PMI and PMI-A determinations has yet to be established.

Table 3-2
Ethanol Contents of the E10 Fuels in the E-94 Studies

Study	E10 Fuel	SwRI ASTM DHA (E-94-2/3) vol%	SSI AVFL-29 DHA (This Study) vol%	Relative Percent Difference
E-94-2	Fuel A	9.04	8.86	-2%
	Fuel B	9.08	9.08	0%
	Fuel E	9.31	9.24	-1%
	Fuel F	9.28	9.01	-3%
E-94-3	Fuel C-E10	9.16	9.22	+1%
	Fuel D-E10	9.41	9.06	-4%
	Fuel G-E10	9.54	9.34	-2%
	Fuel H-E10	9.57	9.40	-2%
Source: Ethanol contents in this table are taken from unpublished DHAs for the fuels and will differ from the published values in Tables 2-2 and 2-3 which are based on ASTM D4815.				

3.3 PEI

PEI—the Particulate Evaluation Index—is one of two index methods examined in Chapman 2016. As described there, the method exploits the observation that, for typical gasolines, the aromatic hydrocarbons are responsible for most of the PMI (typically 95 percent or more). Thus, speciation of only the aromatic compounds should capture most, if not all, of the predictive power for emissions and will permit methods other than the ASTM DHA to be used. In particular, the authors noted that PIONA can be used for the speciation of hydrocarbons by carbon number needed for PEI.

The PEI index was developed in conjunction with the PASCE index discussed later. The work was performed using a small dataset of six gasoline fuels, both E0 (neat) and E>0 fuels (ones containing ethanol). All of the fuels were finished gasolines and included both certification fuels and fuel blends. Overall, the dataset had PMI values ranging from a low of 0.65 to a high of 1.56, which spans the low to middle range of PMI values in the three emissions studies examined here and in US gasoline[11].

[11] See Section 3.1.1 of the CRC E-94-2 Report (2017) for information on the distribution of PMI values in US gasoline.

The authors describe PEI as follows: "The calculation [of PEI] involves the use of a regression equation based on the aromatics from the fuel, and has been derived based on various fuel data sets in comparison to the PMI number." (Chapman 2016, p. 2). The form of the regression equation and its empirical coefficients were provided by the authors for use in this study (Chapman 2018).

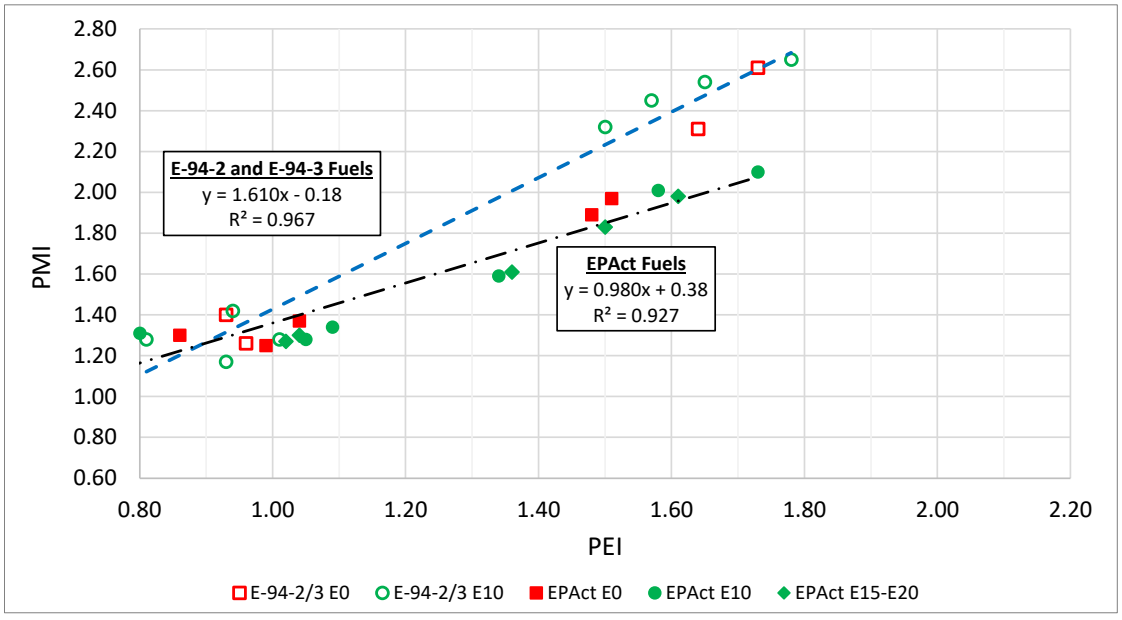
The PEI equation is based on the volume percentages of C7 through C13 aromatic hydrocarbons in the fuel, expressed as decimal fractions, with empirical coefficients for each group determined through regression analysis with PMI. The index calculation is given by:

$$PEI = C7_{AROM} + 2.5 \cdot C8_{AROM} + 5.8 \cdot C9_{AROM} + 4.8 \cdot C10_{AROM} + 35.3 \cdot (C11_{AROM} + C12_{AROM} + C13_{AROM}) \quad \text{(Eq. 3-2)}$$

Here, the variables $C7_{AROM}$ through $C13_{AROM}$ are volume percents (as decimal values) for the aromatic hydrocarbons present in the fuel. The equation makes clear the importance of heavy aromatic hydrocarbons from the 35.3 coefficient for the C11-C13 aromatics group. According to the index, this group exerts 35.3 times the weight of the C7 aromatics in determining a fuel's potential to form PM. PEI is a speciation-based index method that employs empirical coefficients, rather than the "i-term" of the Honda Equation, to estimate the contribution of the carbon numbers to a fuel's PM potential.

Figure 3-2 shows the correlation of PEI to the PMI. For the E-94 fuels (upper blue line), the index correlates strongly (R^2 = 0.967) with PMI. Eight of the 12 fuels were match-blended in E-94-2, while four additional fuels were splash-blended in E-94-3 from the E0 fuels used in E-94-2. The strong correlation means that PEI should perform well in predicting emissions for these fuels.

Figure 3-2
PEI Correlation to PMI



The correlation of the indices for the EPAct fuels (lower black line) is also good (R^2 = 0.927), but not as strong as for the E-94 fuels. The different slopes in the figure for the E-94 studies and for EPAct indicate that different sets of empirical constants would be needed in the PEI equation to give the best match to PMI for their fuels. The EPAct fuels were created to match the maximum range of bulk fuel properties for commercial gasoline and included some fuels with atypical combinations of some properties. As a result, the EPAct fuels are likely to be more dissimilar to the fuels used in developing PEI than the fuels in the two E-92 studies. The correlation of PEI with the PMI, used here as the gold standard for comparison, may be adversely affected by this and also by the less-detailed hydrocarbon speciation it uses.

3.4 Inverse Smoke Point and Oxygen Extended Sooting Index

Two of the indices examined in this study—Inverse Smoke Point (1/SP) and Oxygen Extended Sooting Index (OESI)—are taken from Barrientos 2016, which explored the use of smoke point measurements and other fuel characteristics to define indices for the sooting tendency of gasoline fuels in SIDI engines. As a starting point, the work used smoke point (SP) measurements in the form 1/SP as an index for PN and PM emissions. Then, it introduced terms for additional fuel characteristics (molecular weight and variables representing fuel molecular structure such as branching, degree of saturation, etc.) to define and examine a sequence of indices. The work used existing sets of fuels and emissions testing (e.g., from Aikawa 2010) to demonstrate the correlation of particulate indices to PN and PM emissions; no new emissions testing was performed.

Barrientos 2016 considered two sets of fuels created using a US Tier 2 emissions certification gasoline as the base fuel. The first set involved the Aikawa 2010 experimental fuels created by blending 7 individual hydrocarbons at the 10 wt% level and ethanol at the E21 level for one fuel. The correlation coefficients of 1/SP and OESI with PN emissions were 0.92 and 0.91, respectively (vs. 0.94 for PMI), with the E21 fuel being furthest from the trend line. The second set created oxygenated fuels by blending ethanol, n-Butanol and Isobutanol at varying levels from 10% by volume up to 100%. The oxygen content range in these fuels extends far beyond the E10 to E20 levels found in the three emissions studies, making this part of the experimental work not comparable to the studies considered here. Only the base fuel is comparable to finished gasolines in the US market.

For the work performed here, four different particulate indices were given preliminary consideration in an initial comparison of index performance:

- 1/SP
- TSI^*
- $OESI^*$-DHA MW
- $OESI^*$-Est MW

The * superscript indicates that unscaled smoke point values were used in the calculations, rather than scaled values that would permit comparisons across laboratories. Ultimately,

the 1/SP and OESI*-DHA MW indices were selected for inclusion in this study because they span the range of index complexity and their performance in predicting emissions was found to span the performance range for the group.

The first index is the mathematical inverse of the measured smoke point, written as 1/SP. As noted by Barrientos 2016, smoke point has long been used to characterize the sooting tendency of diesel and aviation fuels. The other three indices—TSI*, OESI*-DHA MW and OESI*-Est MW—are derivatives of 1/SP created by adding additional terms in an attempt to better describe the sooting tendencies.

The threshold sooting index (TSI or TSI*) defined by Calcote 1983 is the second index considered here. In its full form, TSI is given by:

$$TSI = a \bullet (MW / SP) + b \qquad \text{(Eq. 3-3)}$$

where MW denotes the molecular weight of the gasoline and a and b are experimental constants. The two constants are specific to the smoke point measurements and permit placing measurements by different laboratories on the same scale.

The smoke point values for the E-94 fuels were determined using the ASTM D1322 procedure[12] with smoke point values measured in mm, but the (a, b) constants for scaling between test labs were not determined. Therefore, the TSI index used in this study is given by:

$$TSI^* = MW / SP \qquad \text{(Eq. 3-4)}$$

where the * superscript denotes the use of unscaled SP values (effectively assuming $a = 1$ and $b = 0$). Note that the specific values of a and b have no effect in a linear correlation of TSI with PM emissions values, but could affect non-linear correlations. MW was determined from the hydrocarbon speciation given by the ASTM DHAs, which were available for the E-94 fuels and provide the most accurate method for MW determination. In real world applications of the TSI* index, one would expect to determine MW by a more approximate method, such as one of several available statistical correlations. Using 1/SP instead of TSI* effectively assumes that the fuels being considered have indistinguishable MWs.

Previous work by the authors (see Barrientos 2013) examined OESI*, which was found to provide an improved representation for oxygenated fuels by normalizing the TSI index for the molar stoichiometric oxygen-fuel ratio. OESI* can be written in several mathematical forms. The form used in this study is that defined by Eq. 9 in Barrientos 2016 and given below:

[12] The D1322 method for smoke point applies to kerosene and aviation turbine fuels. Motor gasoline is outside its scope. The measured smoke point values of the E-94 fuels are taken as-is and used here without regard to the applicability and precision of the test method.

$$OESI^{*} = \left(\frac{MW_{Fuel}}{SP}\right) * \left(\frac{Wt_C}{12.0112} + \frac{Wt_H}{4\ x\ 1.008} + \frac{Wt_O}{2\ x\ 15.999}\right) \quad \text{(Eq. 3-5)}$$

In this, the * superscript denotes the use of unscaled SP values, MW is the molecular weight, and the Wt_C, Wt_H, and Wt_O values are the weight percentages of carbon, hydrogen, and oxygen in the fuel (as decimal fractions).

Two different methods were used to determine the molecular weight leading to the alternative indices:

1. For OESI*-Est MW, a MW approximation proposed in Barrientos 2016 is used in which the MW of the base hydrocarbon gasoline is assumed to be 100 g/mole, which is then adjusted appropriately based on the oxygenate content.

2. For OESI*-DHA MW, the actual MW of the fuel as determined from the ASTM DHA is used.

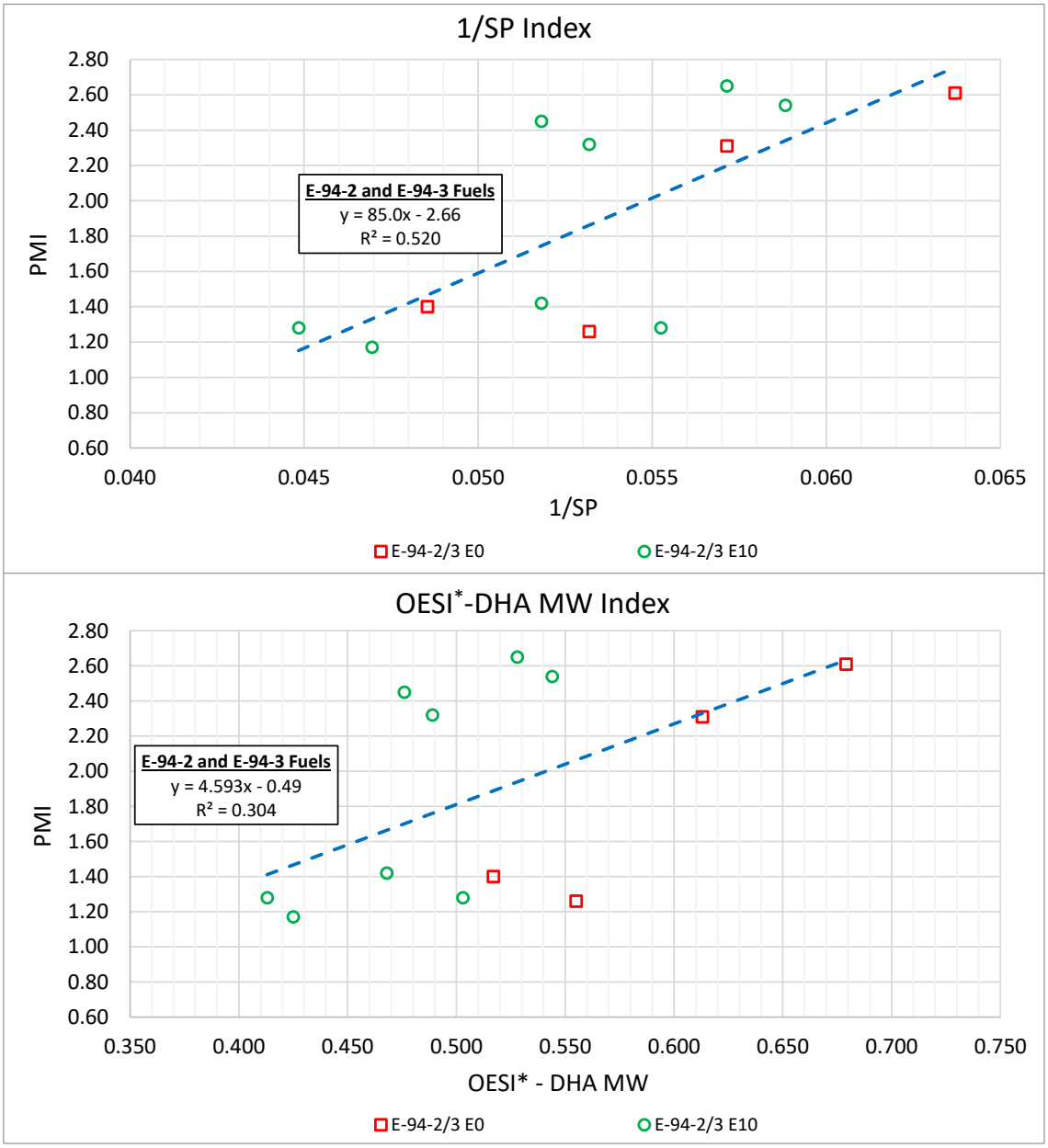
The Barrientos 2016 approximation is a reasonable one given the limited MW range found in commercial E0 gasolines (95-110 g/mole). It is an excellent one for the E-94 E0 fuels, for which the MW ranges only from 98 to 101 g/mole. However, use of the approximation means that only the effect of oxygenation on MW is accounted for while the variation in MW among the gasoline hydrocarbons is ignored. This is a limitation that is expected to degrade performance of the index to some degree compared to determining the MW from the DHA.

Smoke point measurements were not made during the E-94 or EPAct studies. To support this study, SwRI measured the smoke points of the E-94 fuels using retained samples. No retained samples were available for the EPAct fuels, so that it was not possible to determine the 1/SP and OESI*-DHA MW indices. Thus, the EPAct study is excluded from the performance analysis for these indices.

Figure 3-3 shows the correlation of both indices to the PMI using the E-94 fuels. Unlike the other indices considered here, neither 1/SP nor OESI*-DHA MW have strong correlations to the PMI. For 1/SP (top portion), the R^2 value is only 0.520. For OESI*-DHA MW (lower portion), the R^2 value is even smaller at 0.304. For this mixed group of E0 and E10 gasolines, the measured smoke point is not a good predictor of the PMI overall and some data points deviate substantially from the correlation line. There is no particular pattern to the scatter except that it occurs primarily in the middle part of the index range. These fuels can share intermediate values for the 1/SP index but be ranked by the PMI as either low or high PM fuels.

The OESI*-DHA MW index modifies 1/SP with terms for the MW and {C, H, O} ratios with the effect of moving some data points to new locations (left or right) and increasing the scatter around the correlation line. The R^2 value declines as a result. MW has a small range of variation among the E0 fuels and is primarily influenced by dilution of the gasoline hydrocarbons in the E10 fuels; the same is true for the {C, H, O} weight ratios. These factors do not improve index performance for the E-94 fuels considered here.

Figure 3-3
Correlation of 1/SP and OESI*-DHA MW to PMI
(E-94-2 and E-94-3 Studies)



These two indices are distinctive from the others considered here. Not only do they use smoke point as a primary predictor of emissions, but their role in the reference paper is largely as a starting point for further index development, rather than as an index proposed for use with commercial gasolines. The ultimate goal in the work was to develop indices that predicted emissions across the wide oxygenation range in the second dataset (up to 100% oxygenate). Alternative indices considered later in the paper combine the OESI*

index with adjustments for heat of vaporization and net heating value. The 1/SP and OESI*-DHA MW indices may have more promise for further index development than is demonstrated here for the dissimilar group of E0-E10 gasolines in the two E-94 studies.

3.5 PASCE

The Particulate and Soot Correlation Equation (PASCE) is a correlation method defined in Chapman 2016 based on two bulk fuel properties – the C/H ratio (ASTM 5291) and the E170 point (ISO 3405). The correlation of these variables to PMI was determined using a small dataset of six gasoline fuels, both E0 fuels and E>0 fuels. All of the fuels were finished gasolines and included both certification fuels and fuel blends. Overall, the dataset had PMI values ranging from 0.65 to 1.56, which spans the low to middle range of PMI values in the three emissions studies examined here.

The PASCE index is defined by the best fit correlation to PMI for these fuels, given by:

$$PMI = 9.37 + 0.517 * C/H - 0.11994 * E170 \quad \text{(Eq. 3-6)}$$

$$R^2 = 0.983$$

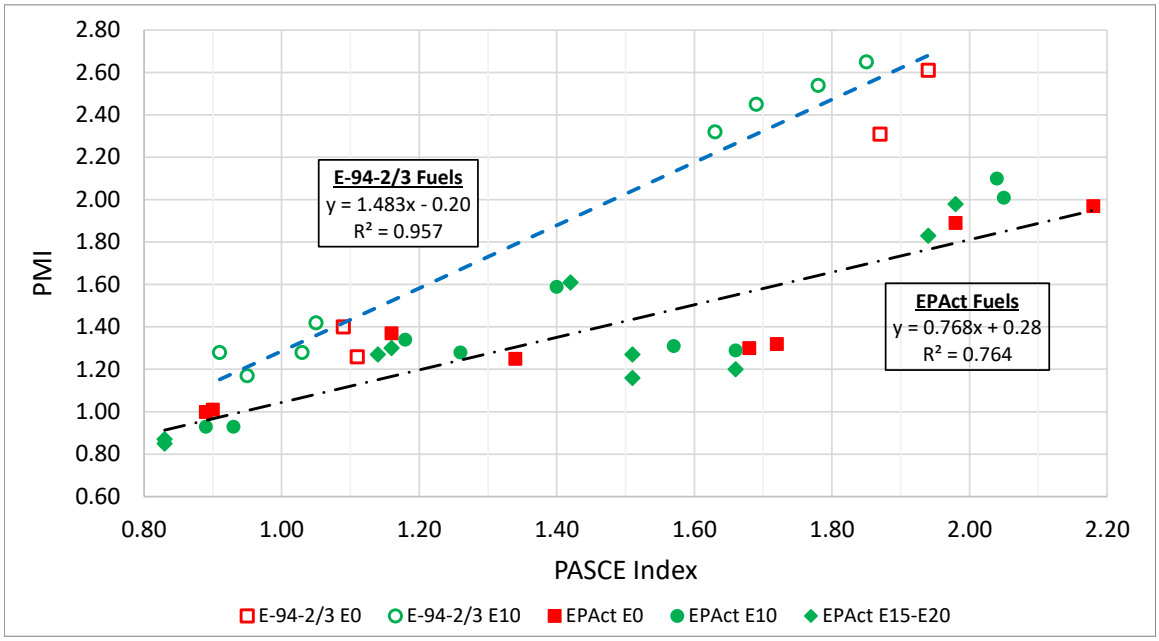
Its authors describe the PASCE correlation as follows. "The correlation equation makes sense from a physical perspective, since diffusion combustion would be a function of the vaporization of the fuel and the specific amount of carbon in the fuel. It also makes sense that this should provide a good correlation to the PMI number, since those values are determined from the fuel vaporization temperature at 443 K (170 degrees C)." (Chapman 2016, p. 2). Correlations with other distillation points in the >100°C range were tested, but the correlations using the E170 point were the strongest.

The PASCE index can be evaluated for the 12 fuels in the two E-94 studies and the 27 fuels in the EPAct study. The C/H ratio was computed from the reported carbon and hydrogen contents (wt%) of the fuels. The E170 distillation point was computed using a spline interpolation of the three ASTM D86 distillation points closest to 170°C. The spline interpolation is sensitive to the local curvature and improves on a simple linear interpolation.

Figure 3-4 shows the correlation of the PASCE index to the PMI for the E-94 fuels (upper blue line) and the EPAct fuels (lower black line). The PASCE index correlates strongly ($R^2 = 0.957$) with PMI for the E-94 fuels. Eight of the 12 fuels were blended to achieve commercial specifications for gasoline in E-94-2, while four additional fuels were splash-blended in E-94-3 from the E0 fuels in E-94-2. The strong correlation means that PASCE should perform well in predicting emissions from these fuels, but it can be seen that the correlation is best for the group of low PMI fuels (at the left). There is more scatter around the correlation line for the group of high PMI fuels (at the right) where we also see that all E10 fuels lie above the line while both E0 fuels lie below. Its performance for high PMI fuels and in representing emissions from both E0 and E10 fuels could fall behind in comparison to the PMI.

For the EPAct fuels, the correlation of the indices is generally good (R^2 = 0.764), but not as strong as for the E-94 fuels. There is also a greater degree of variability around the correlation line within each of the E0, E10 and E15-20 groups of fuels.

Figure 3-4
PASCE Correlation to PMI



3.6 E150 and E130-170 Indices

The E150 and E130-170 indices are simplified indices developed in Moriya 2016 as alternatives to the PMI index method. These indices are the only ones for which a conference presentation, rather than a journal publication, was the best available documentation. After review of the mechanisms of soot formation in gasoline engines, Moriya 2016 says that a series of one and two parameter correlation equations were tested for their ability to predict PN emissions of a test vehicle. The bulk fuel properties considered were: density, a sequence of points on the distillation curve from E130 to E200, RVP, and the total aromatics and olefins contents (wt%).

The PN emissions data were developed from testing of a 2.0L turbocharged direct injection vehicle over the harmonized Worldwide Light-duty Test Cycle using 16 different gasoline fuels. Moriya 2016 does not describe the characteristics of the test fuels or the type and level of oxygenates that were blended, but graphical presentations in the presentation indicate that the PMI values of the fuels ranged from approximately 1.0 to 3.0. This range is somewhat wider than the PMI range of 1.0 to 2.6 in the E-94 and EPAct fuels.

Thirteen different bulk fuel properties were screened for their ability to predict PM and PN emissions using the correlation coefficient. Of these, only the E130 through E170 distillation points exceeded a threshold criterion of $|r| > 0.80$ for further consideration. E150 was the best one parameter correlation with emissions. E130 and E170 were selected from the best two parameter correlation because they showed the lowest pair-wise correlation between the distillation points and, thus, should carry the most independent information about fuels.

Once selected, the one and two parameter correlation equations were used to predict the PMI values of the test fuels. That is, the PMI values were treated as the "gold standard" for comparison and the two correlation equations were tuned to best predict PMI. The E130-150 index, defined by the best fit correlation to PMI, is given by:

$$PMI = 9.9241 - 0.0324 * E130 - 0.0647 * E170 \qquad \text{(Eq. 3-7)}$$

$$R^2 = 0.861$$

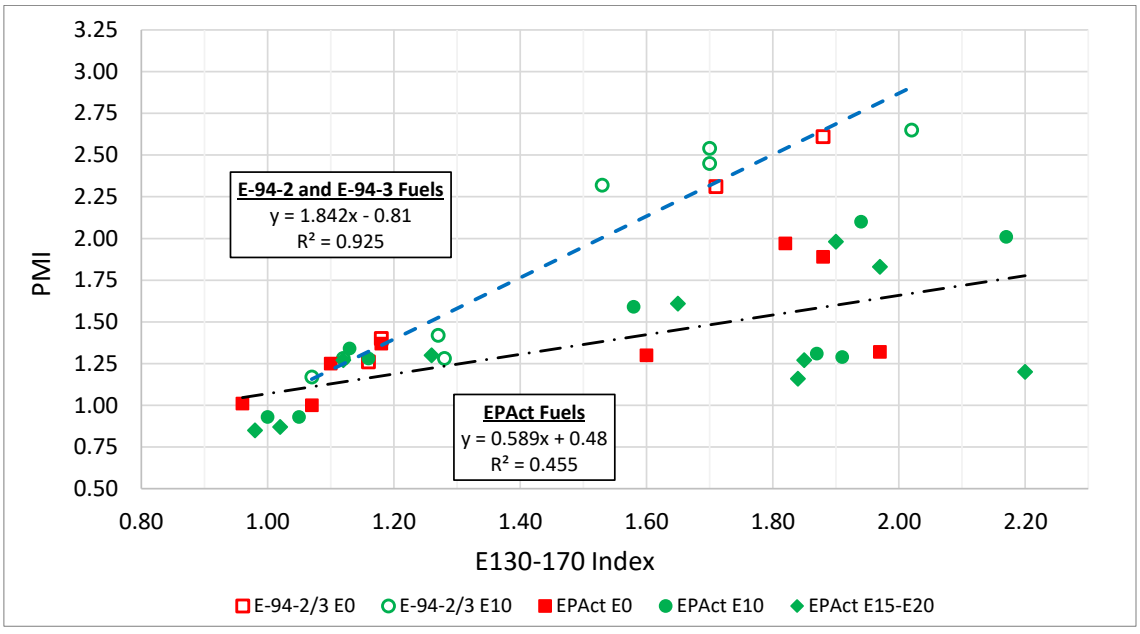
The E150 index is given by:

$$PMI = 7.8511 - 0.0757 * E150 \qquad \text{(Eq. 3-8)}$$

$$R^2 = 0.814$$

Figure 3-5 shows the correlation of the E130-170 index to the PMI for the E-94 fuels (upper blue line) and the EPAct fuels (lower black line). The E130-170 index correlates well ($R^2 = 0.925$) to PMI for the E-94 fuels, but much less well ($R^2 = 0.455$) for the EPAct fuels.

Figure 3-5
E130-170 Correlation to the PMI



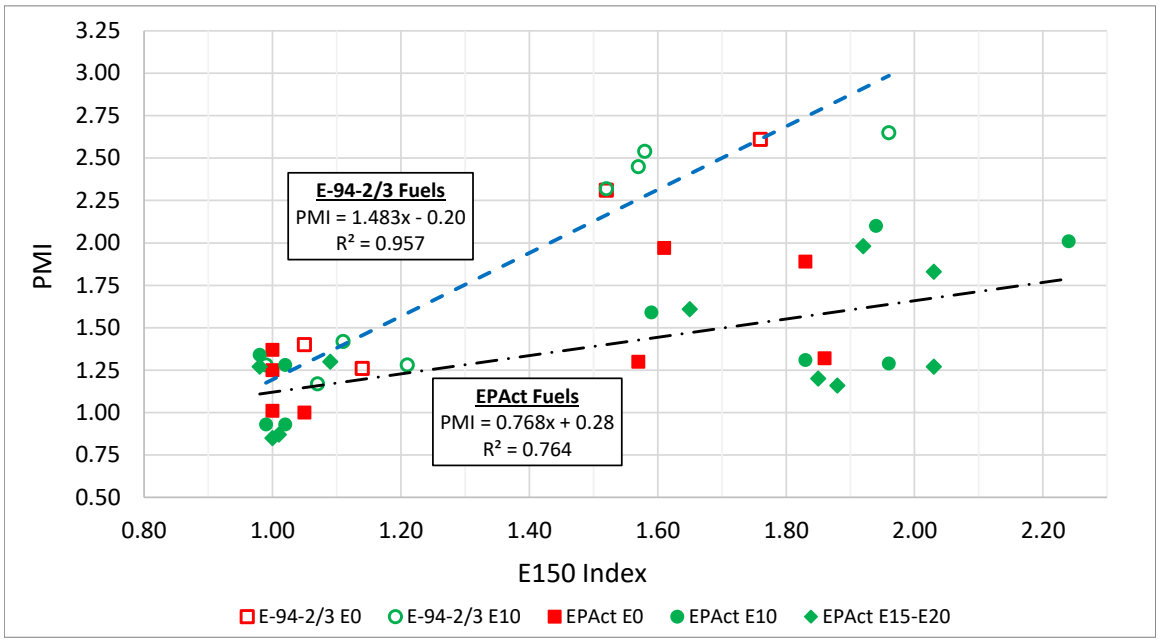
There is no clear-cut pattern in the scatter of data points around the correlation line for the E-94 fuels. For the EPAct fuels, the data points for the high PM fuels (those with E130-170 index values above 1.80) diverge from the correlation line by considerable amounts; in fact, there are no data points on or close to the correlation line. While the E130-170 index ranks all of these fuels as high in PM potential, the PMI divides them into groups: six low PM fuels with PMI values below 1.30 and six high PM fuels with PMI values at or above 1.80. The PMI is able to exploit its detailed information about hydrocarbon composition to draw a distinction between the low and high PM potential of the fuels that the E130-170 index cannot draw, because these fuels are otherwise similar in terms of their E130 and E170 distillation points. Later sections will show that the PMI provides better predictions of PM emissions for the EPAct fuels than either the E130-170 or its sister index E150.

Figure 3-6 presents a similar graph for the correlation of the E150 index to the PMI for the E-94 the EPAct fuels. The E150 index also correlates well ($R^2 = 0.957$) for the E-94 fuels, but less well ($R^2 = 0.764$) for the EPAct fuels. As for its sister index, there is no pattern in the scatter of the E-94 fuels around their correlation line, but for the EPAct fuels the PMI again divides the 12 high PM fuels into two groups of low and high PM fuels in a way that the E150 index cannot do. For the E-94 fuels, the E150 index correlates somewhat better with the PMI than does E130-170, although the difference in the R^2 statistics is not large enough to be material.

For the EPAct fuels, there is a large difference in the R^2 statistics with the E150 index outscoring its sister index by a large margin. The apparent improvement (from R^2 of 0.455 to 0.764) is actually the result of a single E15 fuel (EPAct fuel 27), which is ranked at an index value of 2.20 by the E130-170 index but at a value of 1.85 by the E150 index. The change in index value positions the fuel much closer to the correlation line and substantially reduces its contribution to the prediction error, thereby raising the R^2 statistic. The change moves the fuel from the upper end of the high PM range (E130-170 index) to the lower end of the high PM range (E150 index). In contrast, the PMI value of 1.20 classifies this fuel as a low PM fuel that is at the low end of the PMI range considered in the two E-94 studies.

For E130-170 and E150, it is clear that simple indices can perform well in some cases (the E-94 fuels), but are likely to perform much less well in other cases (the EPAct fuels) where the nature of the fuels differs. The EPAct fuels were designed to span the range of values for five bulk properties: aromatics and ethanol content, RVP, and the T50 and T90 boiling points. In doing so, high aromatic fuels (nominally 35 vol%) were created that spanned a range of T50 temperatures from a low of 150°F to a high of 240°F; additional low aromatic fuels (nominally 15%) were created that spanned the same range in T50. The E150 index is forced to classify these fuels as low or high PM based only on E150 and without knowledge of the aromatics content. The E130-170 index is forced to make a similar classification but has two distillation points to use.

Figure 3-6
E150 Correlation to the PMI



The predicted PM potential of fuels with varying T50 temperatures depends on the total aromatics content and the distribution of aromatic molecules in the fuels. These are unknown to the E130-170 and E150 indices (and to other correlation-based indices), but are accounted for in the speciation-based indices. Thus, it should be no surprise to discover that the E130-170 and E150 indices, along with the other correlation-based indices, may fall behind the performance level of the speciation-based indices when applied to the EPAct fuels.

These results demonstrate two important points about the performance of correlation-based indices and the EPAct fuels examined here:

- Correlation-based indices are "tuned" in the process of the development to the correlations between bulk fuel properties and PM emissions through the empirical coefficients estimated for the chosen properties. The bulk properties reflect and act as surrogates for the hydrocarbon compositions of the test fuels. Whenever such indices are applied to fuels with different relationships between bulk properties and composition, they cannot perform as well. Indices based on hydrocarbon-speciation do not share the same shortcoming because the compositional information is explicitly represented.

- The blending of EPAct fuels to meet specific property targets in a complex experimental design caused the resulting test fuels to depart, to at least some extent, from the correlations between hydrocarbon composition and bulk properties commonly found in gasoline. For commercial gasoline, the total aromatics content

is generally accepted as an indicator of a fuel's potential to form PM[13]. In the EPAct fuels, aromatics content has been decoupled from T50 and other boiling points (and other bulk properties), thereby perturbing the composition-to-property correlations.

Moriya 2016 compares the two simplified indices to the PMI and summarizes their views on the pros and cons of each index using a rating scale of one to four pluses ("+"):

- The PMI and E130-170 have better accuracy for market fuels (+++), while E150 is described as having good accuracy (++)

- The PMI has the best accuracy for unique or atypical fuels (e.g., a heavy paraffin fuel), followed by E130-170 (++) and E150 (+).

- In terms of daily handling, meaning the preparation and evaluation of fuels on a daily basis, E150 received the highest rating (++++), closely followed by E130-170 (+++). The PMI received the lowest rating (+) because of the difficulty in conducting the hydrocarbon speciation.

Later findings in this study add support to the first two of these characterizations. E130-170 proves to be a consistently good performer among the correlation-based indices and, in some settings, equals the performance of the PMI for the gasolines in the two E-94 studies. For unusual or atypical fuels, the PMI (and the other two indices based on hydrocarbon speciation) should perform better than E130-170 because they do not depend on correlations between fuel properties and hydrocarbon composition. This will be shown later when the indices are applied to the EPAct fuels.

3.7 How the Indices Respond to the Presence of Ethanol

An important consideration in the later analysis is how well the indices reflect the effect of ethanol content on PM emissions considering both E0 fuels and E>0 fuels. This aspect of index performance depends on how the presence of ethanol influences PM emissions and the PM index values. An ideal PM index would predict the same level of PM emissions regardless of the content of ethanol in the fuel. None of the indices examined here achieve the ideal, but it should be understood that the ethanol effect on PM emissions is not understood and likely varies with engine characteristics and operating conditions.

The indices differ in how the presence of ethanol changes the index value. For the speciation-based indices—PMI, PMI-A and PEI—the index values respond in proportion to the dilution effect of ethanol. That is, the addition of a given volume of ethanol results in ethanol comprising a corresponding weight percent (wt%) and gasoline hydrocarbons comprising the remainder (1-wt%) of the resulting fuel. If ethanol has no propensity to form PM on its own, as might be expected by its DBE=0 value and high vapor pressure

[13] Also well accepted is that the potential for aromatic hydrocarbons to form PM increases with carbon number. This difference is accounted directly in the speciation-based indices.

and from PM emissions studies with very high ethanol contents,[14] then PM emissions (determined solely by the gasoline hydrocarbons) will decline in proportion to the dilution with ethanol. In this case, the speciation-based indices should perform equally well for E0, E10 and mixed groups of fuels. However, the reality is that the presence of ethanol at the E10 level increases Phase 1 and LA92 PM emissions over and above the level that the PMI of the blended fuel would indicate. Thus, index performance can be expected to vary with the amount of ethanol that is present through the E10-E20 range examined here.

The other, correlation-based indices respond to the presence of ethanol in different ways depending on the fuel property variables contained in the indices. The related 1/SP and OESI*-DHA MW indices are influenced by ethanol through its effect on the measured smoke point. The OESI*-DHA MW index is further modified by the effect of ethanol on the MW and {C, H, O} terms. For these terms, the primary effect is through dilution of the gasoline hydrocarbons and reductions in their values. However, it is unlikely that the potential effect of ethanol, beyond simply dilution, is captured by the smoke point measurements.

The other indices use one of more of the following variables: C/H ratio, E130, E150 and E170 distillation points. PASCE uses the C/H ratio and E170. When ethanol is added to a fuel, the C/H ratio will be reduced as ethanol is a saturated compound with a lower C/H ratio. Ethanol boils at approximately 78°C and the distillation temperatures for a fraction of the fuel (more than twice the ethanol content) are reduced by the addition of ethanol. E170 is particularly diagnostic for the heavy hydrocarbons that are responsible for most of a fuel's PM potential. For the fuels considered here, the E170 distillation point moves by only fractions of 1°C when ethanol is blended at 10 vol%. Thus, the PASCE index is modified primarily through its C/H ratio term.

The E130-170 index is unique among the correlation-based indices in using distillation points that are diagnostic for both ethanol content (E130) and for heavier hydrocarbons (E170). It is affected by ethanol blending through the E130 term, which increases when ethanol is added because a larger portion of the resulting fuel will boil at 130°C or below. This index predicts that E10 fuels will have lower PM emissions than E0 fuels of the same E170 distillation point and is the only one to do so. Its E170 term allows the index to account separately for the portion of heavy hydrocarbons in the fuel. Later sections will show that this index performs well in predicting PM emissions of both E0 and E10 fuels, which likely is a result of the specific distillation points chosen.

The E150 index is a simplification of its sister E130-170 index. Because this point lies between the E130 and E170 points that differentiate ethanol content and heavy hydrocarbons, it will be influenced by both characteristics of a fuel: the addition of ethanol will reduce E150 to some extent, as will the addition of heavy hydrocarbons to the fuel. Thus, the E150 index is not able to distinguish these two characteristics of fuels and proves not to perform as well as its sister index.

[14] For example, see Sakai 2018 and Cho 2015.

3.8 Evaluation of the PM Indices

The section will describe how the values of the PM indices were determined. Tables 3-3 and 3-4 at the end of the section document the index values used here.

3.8.1 Index Evaluation

The PMI and the PMI-A index values were calculated by SwRI and SSI, respectively, where the associated DHA analyses where also performed. The SwRI values for the PMI were reported in the fuel property analyses supporting the two E-94 studies and the EPAct study. The SSI values for the PMI-A were reported in the results of the SSI analysis performed to support this study.

The PEI was evaluated using the profile of aromatic hydrocarbons derived from the ASTM DHA performed for the three emission studies. But, as the index authors note, the inputs in real-world applications could be derived from the PIONA procedure (ASTM D8071). If the accuracy of the PIONA results are comparable to that of the DHAs used here, then the performance of the PEI index using PIONA should be comparable to the performance shown here.

The 1/SP and OESI*-DHA MW indices were evaluated using the smoke point values determined for the E-94-2 and E-94-3 fuels by SwRI from analysis of retained samples. The OESI*-DHA MW index requires additional information on MW and the C, H, and O fractions of the fuels. The MW input was computed from the DHAs for the fuels, while the C, H, and O fractions were taken from fuel analyses reported in the two E-94 studies. Barrientos 2016 discusses the use of approximations for the MW input in real-world applications. The DHA-based MW value was used here to obtain the best possible index performance without any loss associated with an approximate MW.

The PASCE, E130-170, and E150 indices are based on the C/H ratio of the fuel and the E130, E150, and E170 distillation points. The C/H ratio was taken from fuel analyses reported for the three emission studies. The three distillation points were computed using a spline interpolation of the distillation curve. The spline interpolation is sensitive to the local curvature and improves on a simple linear interpolation.

3.8.2 Tabulation of Index Values

Tables 3-3 and 3-4 present the PM index values for the experimental fuels in the two E-94 studies and the EPAct study, respectively. The index values were rounded to two or three significant digits, as shown, before being used in the analysis.

Table 3-3
PM Index Values for the E-94 Experimental Fuels

Fuel	PMI-A	PMI	PEI	1/SP	OESI*-DHA MW	PASCE	E150	E130-170
A	1.43	1.42	0.94	0.0518	0.468	1.05	1.11	1.27
B	3.01	2.65	1.78	0.0571	0.528	1.85	1.96	2.02
C	1.32	1.40	0.93	0.0485	0.517	1.09	1.05	1.18
D	2.97	2.61	1.73	0.0637	0.679	1.94	1.76	1.88
E	1.25	1.28	1.01	0.0552	0.503	1.03	1.21	1.28
F	2.83	2.54	1.65	0.0588	0.544	1.78	1.58	1.70
G	1.23	1.26	0.96	0.0532	0.555	1.11	1.14	1.16
H	2.78	2.49	1.64	0.0571	0.613	1.87	1.52	1.71
C-E10	1.30	1.28	0.81	0.0448	0.413	0.91	0.99	1.12
D-E10	2.75	2.45	1.57	0.0518	0.476	1.69	1.57	1.70
G-E10	1.15	1.17	0.93	0.0469	0.425	0.95	1.07	1.07
H-E10	2.46	2.32	1.50	0.0532	0.489	1.63	1.52	1.53

Table 3-4
PM Index Values for the EPAct Experimental Fuels

Fuel Number	Enhanced PMI	PMI	PEI	1/SP	OESI*-DHA MW	PASCE	E130-170	E150
1	–	0.93	0.63	–	–	0.93	1.05	1.02
2	–	1.32	0.75	–	–	1.72	1.97	1.86
3	–	0.93	0.60	–	–	0.89	1.00	0.99
4	–	1.31	0.80	–	–	1.57	1.87	1.83
5	–	1.37	1.04	–	–	1.16	1.18	1.00
6	–	1.29	0.77	–	–	1.66	1.91	1.96
7	–	1.01	0.69	–	–	0.90	0.96	1.00
8	–	1.00	0.66	–	–	0.89	1.07	1.05
9	–	1.97	1.51	–	–	2.18	1.82	1.61
10	–	2.01	1.58	–	–	2.05	2.17	2.24
11	–	1.34	1.09	–	–	1.18	1.13	0.98
12	–	2.10	1.73	–	–	2.04	1.94	1.94
13	–	1.89	1.48	–	–	1.98	1.88	1.83
14	–	1.30	0.86	–	–	1.68	1.60	1.57
15	–	1.25	0.99	–	–	1.34	1.10	1.00
16	–	1.28	1.05	–	–	1.26	1.16	1.02
20	–	0.87	0.59	–	–	0.83	1.02	1.01
21	–	1.30	1.04	–	–	1.16	1.26	1.09
22	–	0.85	0.59	–	–	0.83	0.98	1.00
23	–	1.27	0.79	–	–	1.51	1.85	2.03
24	–	1.16	0.74	–	–	1.51	1.84	1.88
25	–	1.83	1.50	–	–	1.94	1.97	2.03
26	–	1.98	1.61	–	–	1.98	1.90	1.92
27	–	1.20	0.71	–	–	1.66	2.20	1.85
28	–	1.27	1.02	–	–	1.14	1.12	0.98
30	–	1.59	1.34	–	–	1.40	1.58	1.59
31	–	1.61	1.36	–	–	1.42	1.65	1.65

4. STATISTICAL METHODOLOGY

The analysis of index performance uses conventional regression analysis to determine the relationships between the various PM indices and the Phase 1 PM emissions of the vehicles tested in the three studies. As the large majority of PM is formed during the cold-start phase of the LA92 cycle, the analysis of Phase 1 PM emissions was deemed sufficient to test index performance. The analysis was designed to answer two specific questions:

1. How well does each index perform in predicting actual Phase 1 PM emissions of the test fleets?
2. How well do the PMI-A and the alternative PM indices perform in comparison to the PMI index?

The following sections discuss the key elements of the methodology

4.1 Data Selection and Preparation

The data used in the analysis are the Phase 1 PM emissions of the test vehicles and the characteristics of the fuels on which they were tested. The emissions data consist of individual test runs, while the fuels data span a wide range of physical and chemical properties. The data were obtained from CRC and EPA along with reports that describe the test programs. During this study, additional information was obtained for the E-94-2 and E-94-3 fuels through the analysis of retained samples to determine smoke point and to perform the AVFL-29 DHAs.

The three emissions studies tested the datasets for outliers, meaning anomalous emissions measurements that are unlikely to represent valid data. A small number of individual test runs were identified as probable outliers by the studies' authors and were excluded from the analysis performed here. After the exclusions, the remaining test runs for each vehicle and fuel were averaged to yield a single, composite data point representing the vehicle's Phase 1 PM emissions on the given fuel. Because each vehicle was tested on every fuel in the studies, the three emissions datasets are considered to be balanced.

4.2 Correlation Methodology

Conventional linear regression analysis is used to relate the PM index values to the Phase 1 PM emissions of the test fleets. The resulting regression line is termed a "correlation line" to emphasize that it represents an association between variables.

The dependent variable in the analysis is the natural logarithm of Phase 1 PM emissions. Log form regression models are commonly used in emissions studies because emissions

data can span a wide range and tend to be distributed in a log-normal fashion. The logarithm serves to normalize the data distribution and to equalize the distribution of the log errors across the range of the data. A similar approach was followed in the three emissions studies.

Model 1, the primary mathematical model used in the analysis, is given in Eq. 4-1 below. It includes intercept terms to account for the different emission levels of the vehicles along with the PM index variable to represent the effect of fuels:

$$Y_{ik} = \mu + \nu_k + \text{PM Index}_i + \varepsilon_{ik} \qquad \text{(Eq. 4-1)}$$

where: Y_{ik} = ln(emissions) for fuel i and vehicle k

μ = mean emissions for the average vehicle

i = index for fuels ($i = 1, \ldots, N_i$)

k = index for vehicles ($k = 1, \ldots, N_k$)

ν_k represents vehicle-specific intercepts $\sim N(0,\sigma_v)$

ε_{ik} represents the error term $\sim N(0,\sigma)$

Model 2, a variant that drops the PM Index term, was used in calculating the Incremental R^2 statistic as described later.

Statistical software was used to estimate the models for each index and emissions study. Although the estimation process yields coefficient values for ν_k and the PM Index, the coefficients are not of direct interest in the study. Instead, the models were evaluated to predict emissions for each vehicle/fuel combination. The predicted emissions, the 95% confidence limits of the correlation lines, and the residuals from the lines were computed to permit display of the correlations and data and the evaluation of performance metrics.

4.3 Performance Metrics

R^2 is the usual metric for evaluating the quality of a predictive relationship. In the case of linear regression, R^2 is simply the square of the usual correlation coefficient *r*. For this work, an Incremental R^2 statistic was defined to provide a more direct measure of the ability of the PM indices to explain the variation of PM emissions among the fuels. Two additional metrics are also used in the analysis:

- RMS Error – the root-mean-square error of the data points from the regression line. This metric gives a sense of how large the errors may *typically* be in using a PM index.

- MAE (maximum absolute error) – the absolute error of the data point that deviates the most (up or down) from the regression line. This gives a sense of the *largest* error that might occur in using a PM index.

These three different metrics (Incremental R^2, RMS Error, MAE) are tabulated as overall measures of index performance for predicting PM emissions. Statistical uncertainties in the form of 95% confidence bounds were calculated for R^2 and RMS error using the methods described in Appendix D. A corresponding method could not be identified for MAE, which is used here less as a measure of comparative performance for the indices and more as a measure of risk (the worst-case error committed). Thus, the statistical uncertainties were not determined for MAE.

The Incremental R^2 was defined to provide a better measure of the ability of the indices to predict PM emissions. The problem is that the R^2 values of the correlations are all very high—typically greater than 0.95—because the vehicles differ substantially in their average emission levels. As the regression is performed on individual emission data points for each vehicle and fuel, a large part of the total variation in emissions is simply a reflection of the vehicles' differing average emission levels as represented by the v_k term in Eq. 4.1. By removing this contribution from the R^2 calculation, the Incremental R^2 provides a better, more direct measure of index performance for the fuels.

Table 4-1 illustrates the process of calculating the Incremental R^2. The table is based on the usual organization and terminology for ANOVA output from a statistical package. The Model 1 column shows the output obtained for the correlation of PMI to Phase 1 PM emissions for the E-94-2 study using Eq. 4-1. The Corrected Sum of Squares line gives the total sum of squares around the mean emissions value for the data points; when this is divided by $N-N_k$ (the degrees of freedom), one obtains the variance of the log emissions (Y_{ik}). The Model and Error Sum of Squares lines partition the Corrected Sum of Squares into that explained by Model 1 and the unexplained variation of the data around the line, respectively. The PMI correlation achieves an R^2 value of 0.965 in Model 1, but this is mostly the result of the vehicle contribution.

Table 4-1
Example Calculation of the Incremental R^2 Performance Measure

	Model 1	Model 2	Incremental
Model Sum of Squares	77.755 →	67.647 →	10.109
Error Sum of Squares	2.815	12.924	2.815
Corrected Sum of Squares	80.571	80.571	↘ 12.924
R^2 (Model / Corrected SS)	0.965	0.840	0.782

Note: The black arrows highlight the calculation of the Incremental R^2 as explained in the text.

The Model 2 column shows the output obtained for the variant of Model 1 that drops the PM index term from Eq. 4-1 and includes only the terms related to the average emission levels of the vehicles. The reduced model achieves an R^2 value of 0.840, which is most of the explanatory power of the full model. If this contribution is not removed from the total, the R^2 statistics could mislead the reader into believing that the PM indices explain almost all of the variation in emissions.

The Incremental R^2 is calculated from the Model 1 and Model 2 columns as indicated by the black arrows. The Corrected SS that remains once the vehicle contribution is removed is simply the Error SS (12.924) from the Vehicle correlation, which the black line shows is carried over the Corrected SS row in the Incremental column. The Model SS that the PM index actually explains is the difference (10.109) between the Model SS values in the Model 1 and Model 2 columns. The Error SS not explained by the PM index (2.815) is then calculated from the Corrected and Model SS in the Incremental column. This is necessarily the same as the Error SS in the Model 1 column, as that is the same error that remains once vehicle and PM index contributions have been accounted for. As seen in this example, the Incremental R^2 is 0.782, which represents a good level of performance in predicting emissions, but well below that of the full model.

By removing the vehicle contributions from the calculations, the Incremental R^2 statistic provides a direct measure of how well the PM index explains the variation in PM emissions for the average vehicle in the test fleet. The factors contributing to the incremental Error SS are the unavoidable errors in emissions measurement and the contribution, if any, from differential responses of the vehicles to the fuels. Thus, the Incremental R^2 contrasts the ability of the PM index to explain the variation of emissions across the fuels (for the average vehicle) to the total variation in emissions from all sources.

Three different performance measures were used to assure that conclusions regarding index performance were not unduly influenced by the choice of one measure or another. The three measures have different characteristics and look at index performance in different ways.

The Incremental R^2 gives a relative measure of goodness of fit. Its numerator sums the variation of individual data points around the correlation line, while its denominator sums the total variation of data points around the mean value. This statistic is strongly affected by the size of random errors in the data in comparison to the mean value. The PFI vehicles in the EPAct study have much lower Phase 1 PM emission levels than the SIDI vehicles in the two E-94 studies, but the errors in emissions measurement are approximately the same. As no predictive variable can account for random errors in measurement, the Incremental R^2 values are necessarily lower, and often much lower, for the EPAct study than for the two E-94 studies. The Incremental R^2 is not useful for comparing index performance across the three emissions studies, but does provide useful guidance for comparing the relative performance of different indices for the same study.

The RMS Error is a widely used measure of the typical errors encountered in data. The root-mean-square value carries units of mg/mi, the same physical units as the emissions measurements, but the metric can also be expressed as a percentage by dividing by a representative emission level. Here, the median Phase 1 PM emission values are used as the representative level, rather than the arithmetic average, because the median is a more natural measure of the "central tendency" in data that are log-normally distributed.

The RMS Error in mg/mi terms is a useful guide to the performance of the indices across studies. If the interest is in predicting PM emissions for a group of fuels, or for a large number of fuels individually, this performance measure would be the preferred choice.

The MAE gives the worst-case error committed in applying the PM indices to an emissions dataset. This performance measure is a measure of the risk in using a PM index and may be most useful is deciding whether the index is sufficiently accurate for a particular application. An index that has a large MAE might be bypassed in favor of another with a smaller MAE, even if its typical performance (RMS Error) was equally good or better.

4.4 Potential Improvements to the PM Indices

The objective of this study is to assess the predictive performance of the various indices as they have been defined by their authors. Nevertheless, the CRC committee expressed an interest in obtaining information that might help guide improvements to the indices. To accommodate this interest, additional analysis was performed of the residuals for each index to identify fuel property variables that might add explanatory power if included.

This additional analysis was conducted using an automated, stepwise regression procedure in which the index residuals were tested against each of the physical and chemical properties[15] shown in Table 4-2. The EtOH values used for this purpose are determined from the ASTM D4815 procedure, and not from the DHAs, while the other properties were determined using a variety of standard laboratory procedures. The stepwise procedure will add or remove variables from the residuals model based on their statistical significance.

Table 4-2
Physical and Chemical Properties of the Fuels

EtOH EtOH × Low PMI EtOH × High PMI	IBP	T70
AKI	T05	T80
RVP / DVPE	T10	T90
Aromatics	T20	T95
C10 Aromatics†	T30	FBP
Non-C10 Aromatics†	T40	DI Index
Benzene	T50	Gums Washed
Olefins	T60	Density
† Available in the E-94-2 and E-94-3 studies only.		

[15] While three of the PM indices use E values (fractions evaporated) as predictors, rather than T values (distillation temperatures), the E values were not included in the tests of the residuals. The E values were derived from measurements of the distillation temperatures and are an alternative expression of the distillation characteristics of a fuel.

Because a large number of statistical comparisons are being made, a tightened significance level of $p \leq 0.01$ was required to retain a variable. Stepwise regression is a "blunt" tool in that it is driven only by statistical significance. It will sometimes identify surrogate variables, rather than variables that better represent the underlying causal factors. The variable identifications made in this way are not determinative, but can be used as guidance for further work on the indices.

For the fuel sets involving only E0 gasolines, the results of the stepwise regression process provide the desired identification of fuel property variables that might add explanatory power to the indices. These results are reported in Section 5.5.

For the fuel sets containing E10-E20 gasolines, the analysis was conducted in two steps. First, the stepwise regression process was allowed to select among the variables in the table without restraint beginning with EtOH. A primary objective of this was to test whether ethanol content itself was a significant predictor of the residuals. The results of this step are summarized in Table 4-3. For the E-94-2 and E-94-3 studies, EtOH is a strongly significant predictor by itself, indicating that its inclusion in the PM indices would add predictive power for emissions. In some cases, the stepwise regression process ultimately removed EtOH in favor of other variables, but most commonly ones such as T20, T30 and Density that may serve as surrogates for the presence of ethanol. The analysis had difficulty in detecting an EtOH effect in the residuals for the EPAct study when the dataset was restricted to E0-E10 fuels. For the complete E0-E20 dataset, it found EtOH to be a significant predictor in most cases, although it ultimately replaced it with other variables that may serve as surrogates in all but one case.

Table 4-3
Statistical Significance of EtOH as a Predictor for PM Index Residuals

PM Index	E0-E10 Datasets			E0-E20 Dataset
	E-94-2	E-94-3	EPAct	EPAct
PMI-A	$p<0.001$	$p<0.001$	not evaluated	
PMI	$p<0.001^{\dagger}$	$p<0.001$	$p=0.15^{\dagger}$	$p<0.001$
PEI	$p<0.001$	$p<0.001$	$p=0.39^{\dagger}$	$p<0.001^{\dagger}$
1/SP	$p<0.001$	$p<0.001$	not evaluated	
OESI*-DHA MW	$p<0.001$	$p<0.001$	not evaluated	
PASCE	$p<0.001$	$p<0.001$	$p<0.05^{\dagger}$	$p<0.001^{\dagger}$
E130-170	$p<0.001^{\dagger}$	$p<0.001$	$p=0.41^{\dagger}$	$p<0.05^{\dagger}$
E150	$p<0.01^{\dagger}$	$p<0.001$	$p=0.68^{\dagger}$	$p=0.06^{\dagger}$

Note: † indicates that the EtOH term was ultimately removed as a predictor by the stepwise regression process, most often in favor of T20, T30 or Density which may serve as surrogates for the presence of ethanol.

From the results of the first step analysis, it was concluded that EtOH is an important fuel variable that could be added to all of the indices to improve their predictive power for PM emissions. Thus, a second round of analysis was conducted in which EtOH was a required term in the stepwise regression. The objective in this second round was to identify whether

other fuel property variables might improve the indices' predictive power for emissions once an ethanol effect had been accounted for. The results of this second round are reported in Section 6.5.

4.5 Applications for the PM Indices

The PMI is widely used and accepted in research on PM emissions from gasoline engines because it is based on a detailed assessment of gasoline composition and has proven to be a useful predictor in past work. However, the PMI index requires that a sophisticated DHA procedure be performed for the fuels. Not every laboratory has the capability to perform the DHA and an attempt to use DHAs routinely in daily operations would be time consuming and expensive. Thus, many researchers believe that an alternative, simplified index based on bulk fuel properties would be more useful in daily operations such as fuel blending, provided that it accurately reflects PM emissions effects for the fuel reformulations being considered.

A basic use for PM indices in fuel blending that would be to guide reformulation to reduce PM emissions. For example, if it is desired to reduce the PM potential of a fuel by 10%, one might reblend it in such a way as to reduce its PM index by 10%[16]. Many of the indices considered here would perform well in this application. Which one would be preferred in a given circumstance would depend on a lab's capability for fuel analysis and need for accuracy. This application—the reblending of fuels to achieve a specified emission reduction—is the simplest one to fulfill.

Most of the gasoline sold in the US market contains ethanol at the E10 level. As will be shown in Section 6, most of the PM indices will perform well for fuel groups that have the same ethanol content. It is a different matter when the need is to rank a group of fuels with mixed ethanol levels for emissions—i.e., to put the fuels in order from lowest to highest PM emissions. An equivalent application would be to compare E0 versus E10 gasolines for emissions or to consider adjustments to the ethanol level to achieve an emissions target. None of the PM indices considered here perform well for mixed fuel groups in their current forms.

At issue is the empirical observation that the presence of ethanol in a fuel at the E10-E20 level generally leads to higher PM emissions than would otherwise be expected from the PM index value of the fuel. In mixed sets of fuels, the E10-E20 fuels demonstrate higher emissions than predicted by the correlation, while the E0 fuels demonstrate lower emissions. The PM indices will generally understate the emissions of E10-E20 fuels and overstate the emissions of E0 fuels. If used to guide re-blending for a mixed group of fuels, the indices will lead to an erroneous ranking of the fuels for emissions that tends to favor E>0 over E0 fuels. All of the indices would benefit from additional development work to address this problem.

[16] For the log form of the correlations considered here, PM emissions respond in a non-linear way, but for small emissions changes the relationship is nearly linear.

5. PREDICTIVE PERFORMANCE FOR E0 GASOLINE

This chapter presents the predictive performance of the indices for PM emissions considering only the E0 gasolines in the three studies. This starting point was chosen to establish the ability of the indices to predict the Phase 1 PM emissions of gasoline hydrocarbons before the complicating effects of ethanol are introduced. Many of the indices perform well when ethanol is absent, but some perform better than others. While E10 is the most common fuel in the US market, E0 gasoline is available regionally in the US and it is common in other world markets.

The discussion begins with the PMI, which serves as the "gold standard" for comparisons in this report, and gives a complete presentation of its predictive performance as perspective. Subsequent sections present the comparative performance of the indices and additional information for selected indices to support an understanding of the results. The concluding section summarizes index performance for E0 gasoline and seeks to identify the indices that would be preferred in differing circumstances. Appendix A contains a comprehensive set of E0 results for the indices.

5.1 PMI

The PMI applies the Honda Equation to the hydrocarbon speciation of the ASTM DHA procedure. Figure 5-1 presents the predictive performance of the index graphically. The figure demonstrates the quality of the predictive fit by plotting the data points for each E0 gasoline (in red squares) against the correlation line (in blue) and the 95% confidence interval of the line (in dotted blue). Errors bars (± 1 standard deviation) are given for each of the E0 data points. The upper portion of the figure presents the two E-94 studies, while the lower portion presents the EPAct study. The correlation lines are curved rather than straight because the emissions analysis is performed on the logarithms of emissions (see. Eq. 4-1).

As can be seen, each of the E0 data points overlaps the correlation line within its error bars. This is a visual indication that the Phase 1 PM emissions predicted by the PMI correlation line are not materially different than the actual average PM emissions observed for the E0 fuels. The error bars are relatively small for the two E-94 studies when plotted on a scale suited to the emission values. While the error bars for the EPAct study appear larger in the graph, this is due entirely to the finer scale needed to plot its much lower PM emissions. In fact, the error bars in mg/mi terms are comparable among the three studies. For the three studies and their samples of E0 fuels, the PMI predicts Phase 1 PM emissions to within the errors of measurement for the fuels. This is the standard of performance that the other indices must attempt to match.

Figure 5-1
Predictive Performance of PMI for E0 Gasoline

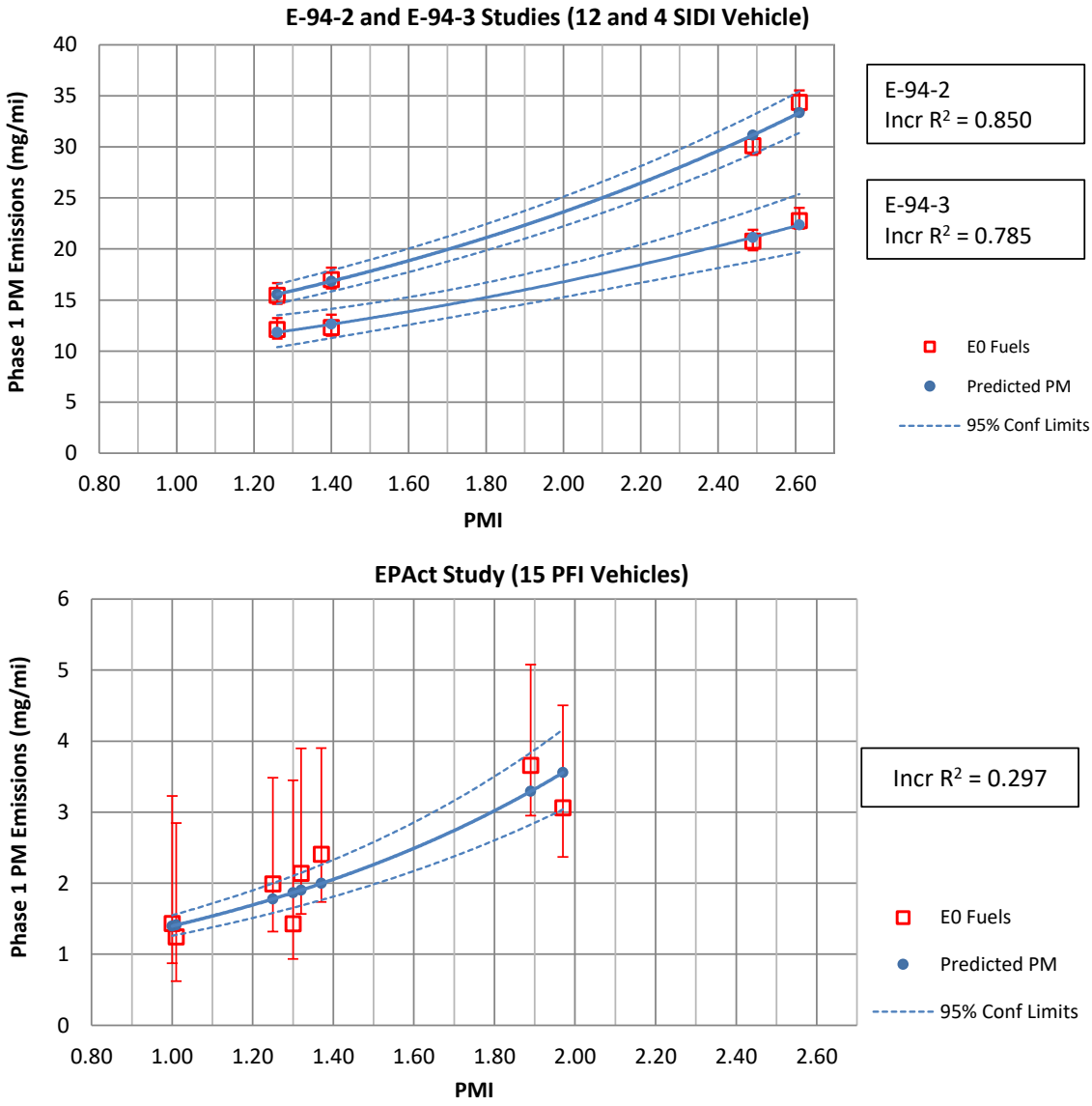


Table 5-1 summarizes the predictive performance of the PMI in the three studies. The first line of the table gives the median Phase 1 PM emissions of the datasets in mg/mi terms; these values are used as the basis to express the RMS and MAE errors in percentage terms. The median value represents the midpoint (50th percentile) of emissions and was chosen, rather than the arithmetic average, as a more natural measure of the central tendency for log-normally distributed emissions data.

The first of the performance metrics is the Incremental R^2, which measures the ability of the PM index to explain the variation of emissions within the data once the effect of the differing average emission levels of the vehicles has been removed. The PMI achieves Incremental R^2 values of 0.850 and 0.785 in the E-94-2 and E-94-3 studies, respectively, and 0.297 in the EPAct study. The value for the EPAct study is substantially lower because

the random errors in emissions measurement play a larger role at the much lower average PM emissions levels of its PFI vehicles. The Incremental R^2 value for the E-94-3 study is marginally lower than the E-94-2 study for the same reason, as its median emission level is marginally lower than the E-94-2. Thus, the differences in Incremental R^2 among the studies do not reflect differences in index performance; rather, they set the baselines against which the other indices will be compared.

Table 5-1
Summary of PMI Performance for E0 Gasoline

	E-94-2	E-94-3	EPAct
Median Phase 1 PM Emissions (mg/mi)	33.0	29.3	2.8
Incremental R^2	0.850	0.785	0.297
RMS Error	2%	2%	16%
mg/mi	0.8	0.7	0.5
MAE Error	3%	3%	31%
mg/mi	1.1	0.8	0.8
PM index of the fuel	2.49	1.40	1.30
Fuel variables adding explanatory power ($p \leq 0.01$)	None	None	AKI

The second and third performance metrics in the table give information on the variation of the data points around the correlation line. The RMS Error, calculated as the root-mean-square deviation of the individual data points, is a measure of the typical error encountered in applying the correlation line to data. On an absolute basis (mg/mi terms), the PMI predicts[17] Phase 1 PM emissions to within ±0.8 and ±0.7 mg/mi for the E-94-2 and -3 studies, respectively, and to within ±0.5 mg/mi for the EPAct study. These values are only 2 to 3 times greater than the precision with which PM emissions can be measured[18] and is a strong indication that the PMI accurately reflects the PM potential for fuels. On a percentage basis, the index predicts PM emissions to within ±2% for the SIDI vehicles in the E-94-2 and -3 studies and to within ±16% for the PFI vehicles in EPAct study.

[17] The term "predicts" is used here and in later sections as shorthand and it should be recognized that the prediction is not solely attributable to the PM index involved. The actual prediction is made using Eq. 4-1 with the PMI as the PM Index variable (in this case) and the empirical coefficients that were estimated for the index by regression analysis. The full prediction involves estimates of the average emission levels of the test vehicles in addition to the PM index's prediction of the PM forming tendencies of the fuels. The prediction is contingent on the LA92 drive cycle used in the testing.

[18] The precision of PM emissions measurement was estimated by SwRI to be on the order of 0.25 mg/mi for the E-94-2 study; similar values can be expected for the E-94-3 and EPAct studies.

The MAE error indicates the maximum absolute or worst-case error that is made in predicting Phase 1 PM emissions for the fuels; the sign of the error (plus or minus) is ignored in its calculation to give equal weight to over-prediction versus under-prediction. The MAE values are comparable in size among the studies, at ±1.1 and ±0.8 mg/mi for the E-94-2 and -3 studies, respectively, and ±0.8 mg/mi for the EPAct study. The table reports the PM index value of the fuel with the worst-case error to permit its identification in the graph and in the original data for the studies. On a percentage basis, the worst-case errors are ±3% and ±3% for the E-94-2 and -3 studies and ±31% for the EPAct study when judged against the median emission levels of the datasets.

Lastly, the table reports the result of statistical tests that were performed to identify fuel variables that could be considered to improve the performance of the index. As explained in the methodology discussion, the tests were performed by regressing the residuals from the correlation line against a wide range of available fuel properties using a stepwise technique. The identifications made here are not conclusive determinations, but rather indications that can provide guidance for additional work.

For the E-94-2 and -3 studies, the statistical tests find that none of the additional fuel properties would add to the predictive performance of the PMI index. This result is consistent with its high Incremental R^2 values and RMS Errors that are not much larger than the expected precision of PM emissions measurement. For the EPAct study, the test identifies AKI as an additional variable to consider. This could be a vehicle effect, as modern vehicles can adjust engine operation to account for differences in octane, but it could also indicate that additional information is needed about systematic differences in hydrocarbon compounds between low and high AKI fuels. If so, the index performance might be improved by expanding the list of compounds considered in the DHA procedure.

In summary, the PMI proves to be an accurate predictor of PM emissions for the E0 fuels in the three studies. It predicts Phase 1 PM emissions to within ±0.5 to ±0.8 mg/mi on an RMS basis, depending on study (a multiple of 2-3 times the expected precision of PM measurement). Its worst-case errors are about ±1 mg/mi (or 4 times the precision of measurement). It is a highly accurate predictor for the SIDI vehicles (±2% on an RMS basis and ±3% on a MAE basis) given their substantially higher emission levels. Finally, it appears to be a complete or nearly complete representation of the PM potential of E0 fuels that would not benefit from the inclusion of the additional variables considered except, possibly, in the EPAct study.

The predictive performance of the PMI establishes a reference point against which the performance of the other indices will be compared. It is termed the "gold standard" for index performance for this reason.

5.2 Comparative Performance of the Indices

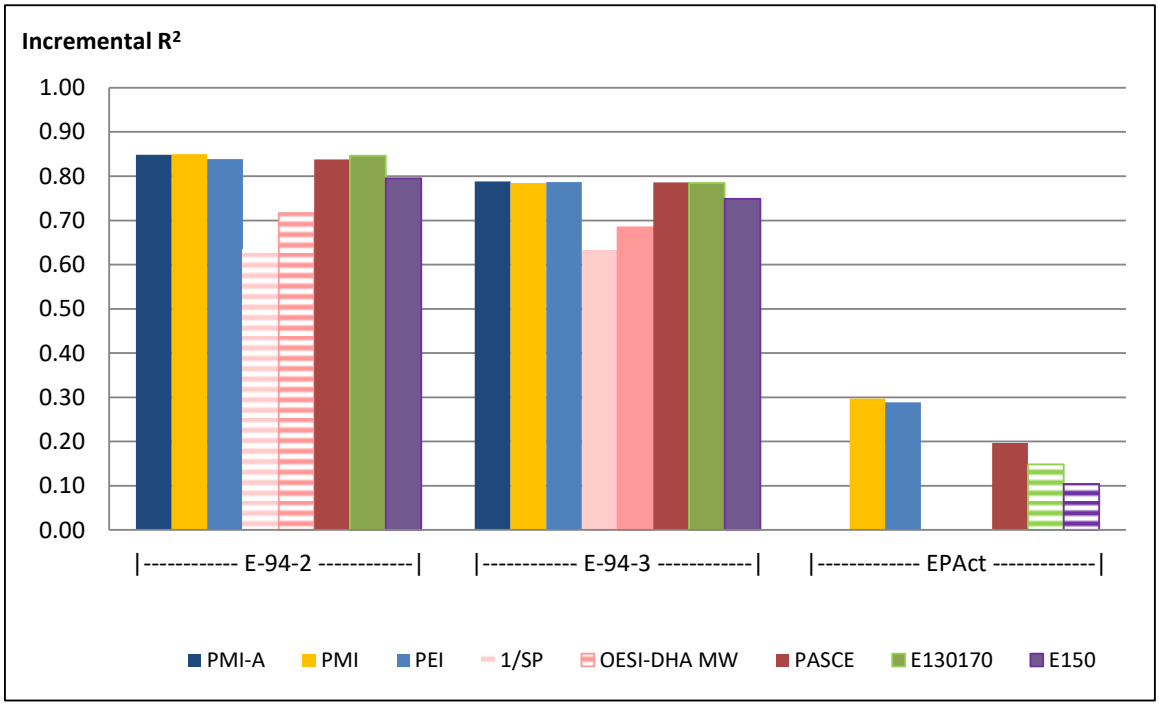
This section examines the comparative performance of the indices based on the three performance metrics. As shown, all of the indices perform well for E0 gasolines and a number of the indices match the performance level of PMI.

5.2.1 Incremental R^2 Performance Measure

Figure 5-2 shows the comparative performance of the PM indices for E0 gasoline based on the Incremental R^2 measure. The indices are presented in groups from left to right for the E-94-2, E-94-3 and EPAct studies. As has been noted, additional fuel analyses were not possible for the EPAct fuels. As a result, the PMI-A and two related indices—1/SP and OESI*-DHA MW—could not be evaluated using the EPAct study.

The indices that are statistically different from the one best performer are indicated by bars with horizontal shading; for them, the difference in performance is enough that one can be 95% confident that they are poorer performers than the one best performing index. For this index measure, the formula for statistical uncertainties cannot be applied to the small sample size of the E-94-3 study. Thus, statistical significance is not evaluated for the E-94-3 study and it is likely that its data are insufficient in number to establish statistically significant differences in index performance.

Figure 5-2
Index Performance based on Incremental R^2: E0 Gasoline



To assess which of the indices are the better performers, the indices have been ranked in descending order of the Incremental R^2 metric, rounded to two decimal places. The indices attaining one of the 3 best scores are selected, including ties, as "high performers" to be specifically noted and discussed. For example, when rounded to 2 significant digits, the top 3 scores for the Incremental R^2 in the E-94-2 study are 0.85, 0.84, and 0.79 and belong to the six indices identified in the text as having high performance. In this, fewer than 3 scores would be included if the indices for a lower score are identified as statistically

different, and poorer, than the one best index. For the EPAct study, only the two best scores are considered because only five of the eight indices could be evaluated.

For the E-94-2 and -3 studies, we see that size of the PM indices demonstrate similar and high performance levels based on the top 3 scores: PMI-A, PMI, PEI, PASCE, E130-170, and E150. For E-94-2, the Incremental R^2 for these five ranges from a low of 0.79 (E150) to a high of 0.85 (PMI and PMI-A). For E-94-3, the Incremental R^2 range from a low of 0.75 (E150) to a high of 0.79 (PMI-A, PEI, PASCE). Such values, differing at most in the second, are the same for all practical purposes.

This result is remarkable, given that three indices differ greatly in terms of the variables involved and the fuel analyses required:

- The Enhanced and PMIs require DHAs performed according to different procedures (AVFL-29 and ASTM, respectively), with the AVFL-29 DHA extending the list of compounds to include heavier hydrocarbons.
- In PEI the treatment of hydrocarbon composition is simplified by: considering only the aromatics compounds; grouping them by carbon number; and weighting their contributions to emissions using empirical coefficients. A speciation of hydrocarbon composition is required, but not to the level of individual compounds as in a DHA.
- The PASCE index is a correlation method based on the C/H ratio and the E170 distillation point.
- The E130-170 index is a correlation method based on the E130 and E170 distillation points, and as is the E150 index based on the E150 distillation point.

The 1/SP and the OESI*-DHA MW indices trail in this comparison due to the reduced performance of smoke point as a predictor of emissions. As shown, later, 1/SP is a good predictor of emissions for two of the E-94 fuels, but not as good for two others. As 1/SP is also a term in OESI*-DHA MW, it pulls down the performance of that index as well. Horizontal bars are used in the figure to mark these indices as being statistically different, and poorer performers, than the one best performer.

For the EPAct study, only PMI and PEI achieve a high performance level (top 2 scores), with the performance measures for PASCE, E130-170, and E150 falling below and the latter two indices being found statistically different from the best. The Incremental R^2 values are 0.30 for PMI and 0.29 for PEI, which is not a meaningful difference. There is a more notable gap between those two and the other indices. If the PMI-A could be evaluated for the EPAct fuels, we would expect it to be a high performer because of the similarity in methodology and the high correlation seen earlier between the PMI and PMI-A for the E-94 fuels.

It has been noted that random error plays a large role in the performance measures for the EPAct study due to the lower average PM emission levels of its PFI vehicles. It is a fact that the EPAct performance levels are lower for all of the indices compared to the E-94-2 and -3 studies, but the result seen here goes beyond the influence of random errors. The

three indices that trail the best performers are correlation methods that rely on bulk fuel properties (C/H ratio, E130, E150 and E170) to predict emissions.

The correlations exploited by the three trailing indices are those between the variables they use as predictors and the hydrocarbon composition of the test fuels that were used to estimate the empirical coefficients in the indices. Because the correlation methods perform as well as the three speciation-based methods for the two E-94 studies, we cab surmise that the correlations in the test fuels from which the indices were estimated are similar to those in the E-94 E0 fuels.

In contrast, the EPAct study design sought to create fuels that spanned the 5^{th} and 95^{th} percentiles of commercial fuels and included some fuels with atypical combinations of properties. This process may lead to hydrocarbon compositions that perturb the correlations of variables such as C/H ratio, E130, E150 and E170 to hydrocarbon composition. The nature of the EPAct fuels design is likely the reason for the lower level of performance for the three correlation-based indices.

This conclusion is important beyond the limits of this study. It illustrates the fact the correlation-based indices can perform well when they are applied to datasets that have similar fuels and correlations. But, correlation-based indices will generally fail to achieve the same level of performance in datasets where the correlations differ; the level of performance that is achieved will vary based on how far the correlations are perturbed from the original dataset. One should expect the same to happen in real-world applications of correlation-based PM indices when they are applied to fuels that have been created using processing and/or blending methods that differ from the practices used in making the fuels on which the correlations were based.

5.2.2 RMS Error Performance Measure

Figure 5-3 shows the comparative performance of the PM indices as indicated by the RMS Error measure. As before, the indices are presented in groups from left to right for the three studies. Here, statistical uncertainties in the performance measures can be determined for all three studies. Columns that are statistically different from the one best performer are marked with horizontal bars. Here, indices are ranked in order of performance based on the RMS error rounded to 2 significant digits (0.1 mg/mi). Indices attaining one of the 3 best scores (2 best for the EPAct study) are discussed as “high performers”.

Overall, the relative performance of the indices is similar to that seen above, but fewer indices are high performers and larger differences are apparent among the indices because of the scale of the graph. For the E-94-2 study, PMI is the best performer with an RMS Error of ±0.8 mg/mi. PMI-A and E130-170 are close second and third best performers, at ±0.9 and ±1.1 mg/mi respectively. All other indices have RMS Errors well above 1 mg/mi and are found to be statistically different (poorer) than the one best performer. The E150 index has a substantially larger error than the E130-170 index. The 1/SP index has the largest RMS Error of all (±6 mg/mi) and, again, degrades the performance of OESI*-DHA MW.

For the E-94-3 study, PEI is the one best performer with an RMS Error of ±0.5 mg/mi followed in order by PASCE at ±0.5 mg/mi, E130-170 and PMI at ±0.7 mg/mi. PMI-A follows close behind the high performers at 0.8 mg/mi, which is the 4th best score. E150, 1/SP and OESI*-DHA MW show much larger errors and are found to be statistically different than the one best performer. Note that the fuels examined here are the same fuels used in the E-94-2 study, with the difference being that only 4 of the 12 SIDI vehicles are considered. Thus, the performance difference seen here between the E-94-2 and E-94-3 studies is related to the subset of vehicles involved and not to the underlying index performance.

For the EPAct study, we see that PMI and PEI are the two high performing indices and out-perform PASCE, E130-170 and E150 by a statistically significant margin. Given the high correlation between PMI and PMI-A, we would expect PMI-A to join them if it could be evaluated for the fuels. The indices based on speciation of hydrocarbon composition have an advantage over indices based on correlation methods for the EPAct fuels.

Taken together, the results point to PMI, PMI-A and PEI as being high performers on a consistent basis. The E130-170 index is a close competitor for the commercial-spec gasolines in the E-94-2 study and PASCE and E130-170 for the E-94-3 gasolines.

5.2.3 MAE Performance Measure

Figure 5-4 shows the comparative performance of the PM indices as indicated by the MAE measure. As noted, the statistical uncertainty was not evaluated for the MAE metric. As for the RMS Error metric, Here, indices are ranked in order of performance based on the RMS error rounded to 2 significant digits (0.1 mg/mi) and ones attaining one of the 3 best scores (2 best for the EPAct study) are discussed as “high performers”.

Overall, the pattern of performance for the indices is very similar to that for the RMS Error. The primary difference is that the numerical values are larger because we are looking at worst case errors. For E-94-2, PMI-A is the one best performer, followed closely by PMI and E130-170. PEI is the one best performer for E-94-3 followed by PASCE and PMI. The Enhanced and PMIs are the clear best performers for the EPAct study and likely would be joined by the PMI-A if it could be evaluated. As was the case for the other performance measures, the indices based on speciation of hydrocarbon composition have an advantage over the indices that are based on correlation methods.

Figure 5-3
Index Performance based on RMS Error: E0 Gasoline

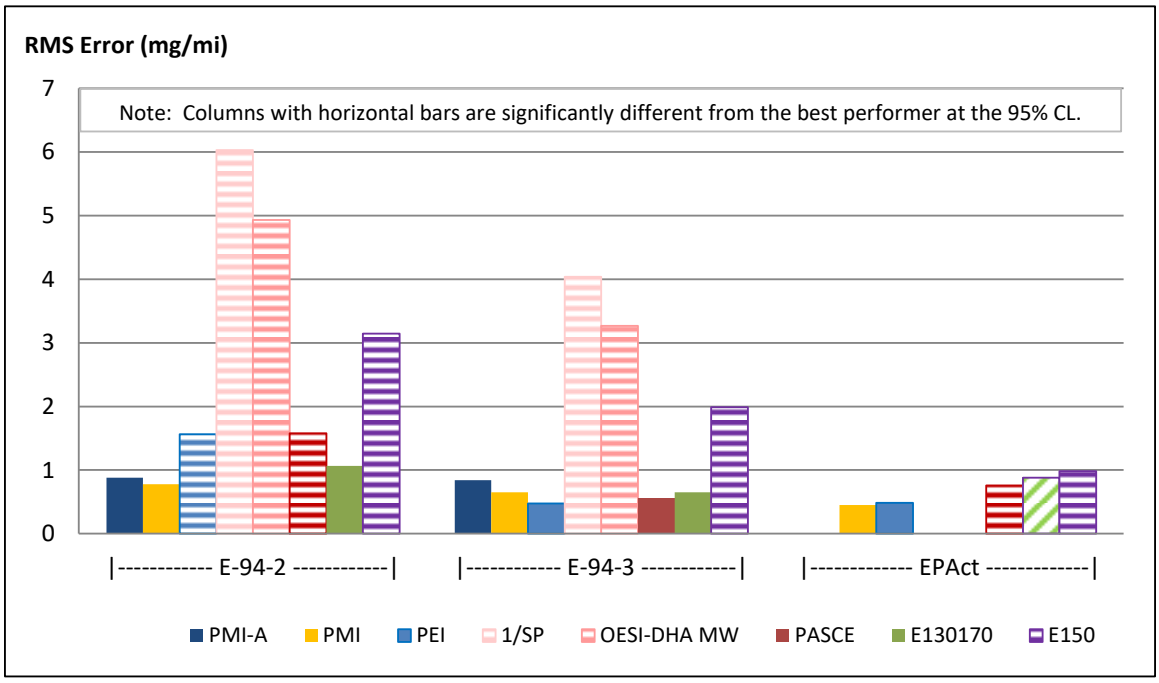
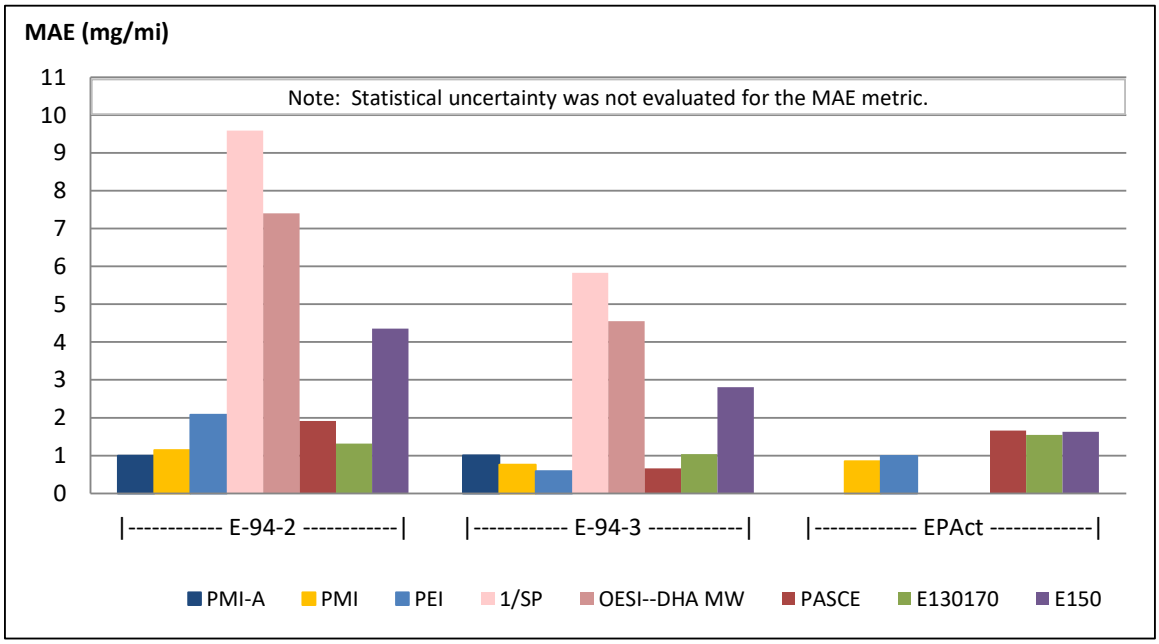


Figure 5-4
Index Performance based on MAE: E0 Gasoline



5.3 PMI versus PMI-A

The AVFL-29 DHA procedure was developed to extend the speciation of the heavier gasoline hydrocarbons beyond that covered in the ASTM DHA procedure. Although the heaviest hydrocarbons might constitute only small portions of fuels, they contribute disproportionately to PM emissions. In the preceding figures, we saw that the PMI-A performed as well as the PMI in predicting Phase 1 PM emissions, but it was not clearly better.

Table 5-2 summarizes the performance measures for PMI and PMI-A. For the E-94-2 study, the PMI has slightly better performance when measured as Incremental R^2 and RMS Error, but the PMI-A is slightly better in terms of MAE. For the E-94-3 study, the PMI has slightly better performance for the RMS Error and MAE metrics, but the PMI-A is slightly better in terms of Incremental R^2. The differences are not large enough to be meaningful in a practical sense and we might see directionally different comparisons if another group of similar fuels were tested. There is no basis here to prefer one index over the other as the comparison is based on a relatively small number of fuels and vehicle tests. More data and further comparisons would be needed to better answer this question.

Table 5-2
Performance Comparison for PMI and PMI-A: E0 Gasoline

	E-94-2 Study		**E-94-3 Study**	
	Original PMI	PMI-A	PMI	PMI-A
Incremental R^2	0.850	0.849	0.785	0.788
RMS Error	2.4%	2.7%	2.2%	2.9%
mg/mi	0.8	0.9	0.7	0.8
Maximum Absolute Error	3.5%	3.0%	2.6%	3.4%
mg/mi	1.1	1.0	0.8	1.0

While systematic differences between the test labs may be present, another factor is likely to be more important in understanding the performance of the indices. PMI and PMI-A are closely correlated variables for the E-94-2 and -3 fuels, as was demonstrated previously by Figure 3-1. When using correlated variables, one variable (PMI-A) might be a more complete and reliable assessment of a characteristic, but it may not demonstrate notably better predictive performance than another variable (PMI) when the correlation between the two is very high. Thus, the AVFL-29 DHA may yield an improved understanding of hydrocarbon composition in heavy gasolines, but it will not necessarily be a better predictor, at least for the fuels considered here.

Gasolines available in the US market are blended from refinery streams to meet commercial specifications, but without specific attention to the composition of heavy hydrocarbons. In this circumstance, and assuming similar crude slates and refining

operations and blending strategies among refineries, the values for the heaviest hydrocarbons that are identified only in an AVFL-29 DHA should be correlated to the values for heavy hydrocarbons that are identified in an ASTM DHA. Thus, for typical fuels, the PMI may be able to predict PM emissions as well as the PMI-A, even if the PMI-A gives a more complete assessment of the hydrocarbon composition.

In other cases, where the prevalence of the heaviest hydrocarbons in a fuel is atypically higher (or lower) than the prevalence of other heavy compounds, the PMI-A should be a better and more reliable predictor of emissions. This can be seen in the performance measures, although only to a very small extent. PMI and PMI-A show essentially equivalent performance based on the Incremental R^2 and RMS Error, but the PMI-A shows slightly better performance based on the MAE (the worst-case error). In an experiment where High PMI fuels were created in different ways, using differing crudes and blendstocks, the PMI-A could well prove to be a superior predictor.

5.4 The 1/SP Indices

The 1/SP index, and the related OESI*-DHA MW index, demonstrate a substantially lower level of performance in predicting Phase 1 PM emissions than the other indices. The reason for this is that 1/SP gives an inconsistent representation of emissions for some fuels, unlike the other indices where the accuracy in predicting emissions is generally similar across the range of PM index values. Because 1/SP is the first term in the OESI*-DHA MW index, its inconsistent performance also affects the latter index and to an extent that is not overcome by its other terms.

Figure 5-5 demonstrates the inconsistency of 1/SP. While Phase 1 PM emissions trend upward as the 1/SP values increase, two of the four data points lie well off the correlation line and at distances that are many times larger than their associated error bars. This is an indication that 1/SP represents some, but not all, of the factors that influence emissions. The 1/SP index can be a poor predictor of emissions for some fuels (such as the middle two in the figure), but a good predictor for other fuels (the leftmost and rightmost fuels). The nature of the information that is missing from 1/SP is not known, but may relate to the evaporative characteristics of the individual or bulk gasoline as suggested by Barrientos 2016.

5.5 How the Indices Might be Improved for E0 Gasoline

The last step in the analysis for each index was to search for additional fuel variables that might improve the performance of the indices for E0 gasoline. This was done using stepwise regression techniques in which the residuals of the index correlations were tested against a list of available fuel properties. Separate tests and determinations were made for each index and study. The variables identified by this approach are not conclusive determinations, but rather indications that may provide guidance for future work.

Figure 5-5
Predictive Performance of Inverse Smoke Point for E0 Gasoline

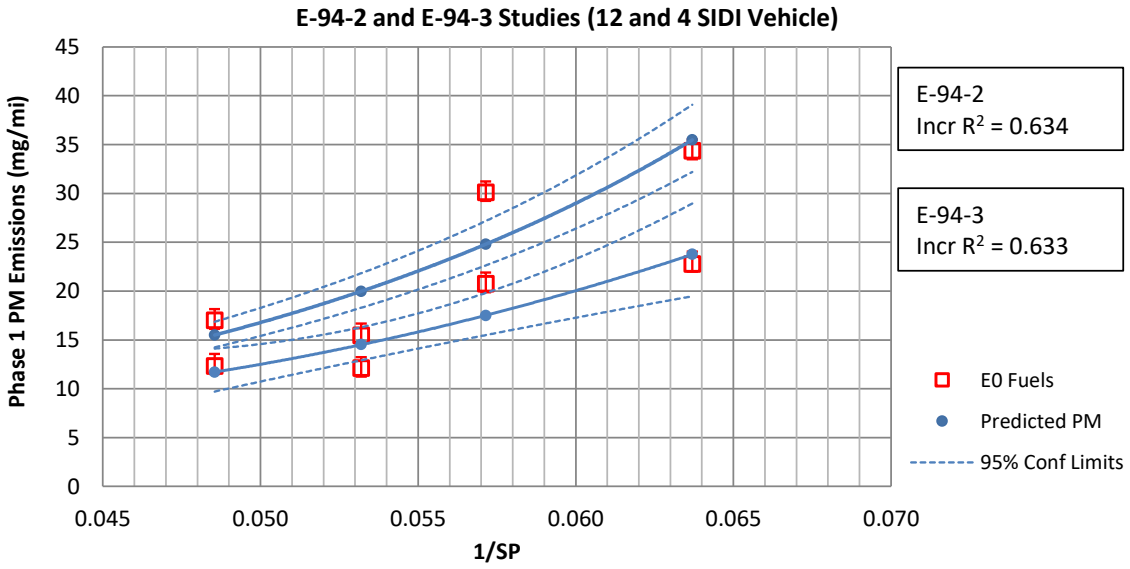


Table 5-3 summarizes the variables that were identified in this process. In the E-94-2 and -3 studies, most of the indices would not benefit from the inclusion of additional fuel variables (indicated as "—" in the table). The 1/SP and OESI*-DHA MW indices, which lagged in predictive performance, would benefit from the inclusion of fuel density to account for the prevalence of aromatic hydrocarbons versus saturated compounds and light versus heavy hydrocarbons. E150, the simplest of the PM indices, also would benefit from including density as a variable, as its use of a single distillation point at E150 cannot account for composition differences to the extent that E130-170 (its sister index) can. In general, density is a logical choice for inclusion in the PM indices because it is related to the prevalence of aromatic hydrocarbons and light versus heavy hydrocarbons in a fuel.

In contrast, all of the indices that were evaluated for the EPAct fuels could benefit from additional variables. AKI was identified as a variable that could improve the predictive performance of the PMI. This could be a vehicle effect, as modern vehicles can adjust engine operation to account for differences in octane, but it could also indicate that additional information is needed about some hydrocarbon compounds in the high AKI fuels. This is a case where the PMI-A might have demonstrated an advantage over the PMI.

For PEI, the predictive performance could be improved by the addition of the T60 temperature on the distillation curve, which could be a surrogate variable for PM contributions from non-aromatic compounds in the fuels. This would be in addition to improvements that might arise from changes in the index formula, such the disaggregation of the C11-C13 category of aromatic hydrocarbons. For the other indices—PASCE, E130-170 and E150—density is again identified as a variable that could improve the predictive performance.

Table 5-3
Fuel Property Variables That Might Improve the Indices

	E-94-2 Study	E-94-3 Study	EPAct Study
PMI-A	—	—	(Not Evaluated)
PMI	—	—	AKI
PEI	—	—	T60
1/SP	Density	Density	(Not evaluated)
OESI*-DHA MW	Density	—	(Not evaluated)
PASCE	—	—	Density
E130-170	—	—	Density
E150	Density	—	Density

The fact that all of the indices could benefit from additional variables when applied to the EPAct fuels is another indication that the EPAct fuels depart, to at least some extent, from the correlations between hydrocarbon composition and fuel properties found in the gasolines used to develop the indices. With respect to PMI and PEI, the departures could indicate the presence of hydrocarbons that the indices do not account for, that the PMI equation does not weight the included compounds perfectly, or that other fuel factors contribute beyond the DBE and VP variables that are accounted for in the equation. With respect to the other indices (PASCE, E130-170 and E150), additional variables may be needed to account for the generally less-typical compositions of the EPAct fuels.

5.6 Summary of Index Performance for E0 Gasoline

Table 5-4 compiles a summary of index performance for E0 gasoline in an effort to find consensus among the performance metrics. For each of the metrics, the indices identified as high performers based on the 3 best scores (2 best for the EPAct study) are indicated in the table by their rankings, ties included.

The three indices based on hydrocarbon speciation—PMI-A, PMI and PEI—are the ones most often classified as high performers. The PMI and PEI are the only indices classified as high performers for the EPAct study, but one would expect the AVFL-29 DHA to have been a high performer as well. Among these, PMI receives the most high performer rankings (9), but is followed closely by PEI (7). PMI-A likely would have tied with its PMI counterpart had it been evaluated for the EPAct study. When one of these indices is not the one best performer, it usually ranks as a close contender. Thus, one's preference for one of these three indices over the other two is more likely to be determined by one's capability and preference for hydrocarbon speciation, rather than the performance of the index itself.

Table 5-4
High Performing Indices for E0 Gasoline

Index	Study	R^2	RMS Error	MAE
PMI-A	E-94-2	#1	#2	#1
	E-94-3	#1		
	EPAct	(Not Evaluated)		
PMI	E-94-2	#1	#1	#2
	E-94-3	#2	#3	#3
	EPAct	#1	#1	#1
PEI	E-94-2	#2		
	E-94-3	#1	#1	#1
	EPAct	#1	#2	#2
1/SP	E-94-2			
	E-94-3			
	EPAct	(Not Evaluated)		
OESI*—DHA MW	E-94-2			
	E-94-3			
	EPAct	(Not Evaluated)		
PASCE	E-94-2	#2		
	E-94-3	#1	#2	#2
	EPAct	#2		
E130-170	E-94-2	#1	#3	#3
	E-94-3	#2	#3	
	EPAct			
E150	E-94-2	#3		
	E-94-3			
	EPAct			

Among the other indices, PASCE and E130-170 contend for consideration as high performers. For the E-94-2 and -3 studies, they are the one best performers for the Incremental R^2 metric, but not for the RMS Error or MAE metrics where they rank as close contenders. The two indices perform well overall, but they are subject to somewhat larger errors in general and, specifically, for some individual fuels. For the EPAct study, neither index is a high performer or a close contender, likely because the design and formulation of the EPAct fuels has perturbed the fuel property correlations away from those on which the indices were based.

The PASCE and E130-170 indices are good performers for the E0 fuels in the two E-94 studies and may also be good performers for other, similar E0 fuels. One would choose

E130-170 if only the distillation curve were known, but either index could be used when both the C/H ratio and the distillation curve are known.

The above conclusions were drawn based on the E-94 and EPAct emissions studies, which provide internally-consistent data for vehicle PM emissions and fuel characterization. The EPAct study was planned and completed before the papers on PMI and other PM indices were published. As a result, analyses of PM index performance using the EPAct data set are retrospective in nature. In addition, it considered only vehicles with PFI engines. In contrast, the E-94 study involving vehicles with SIDI engines, was conducted primarily to focus on PMI and ethanol content as parameters to describe PM effects. In the future, advance planning with multiple indices in mind may strengthen such emissions studies. Additional data from comprehensive vehicle-fuel PM emissions studies would be useful to confirm the conclusions drawn in this report.

6. PREDICTIVE PERFORMANCE FOR E0 AND E10 GASOLINES

This chapter presents the predictive performance of the indices considering only the group of E0 and E10 gasolines in the three studies. This group includes the gasolines that are widely available in the US market, as most contain ethanol at the E10 level while E0 remains available in some areas of the US (and is prevalent in other global markets). The full set of gasolines (E0 through E20) will be examined in the next chapter.

The discussion begins with the PMI, which serves as the "gold standard" for comparisons in this report. Subsequent sections present the comparative performance of the indices based on the three performance metrics. Additional information is given for selected indices to support an understanding of the results. The concluding section summarizes index performance for E0 gasoline.

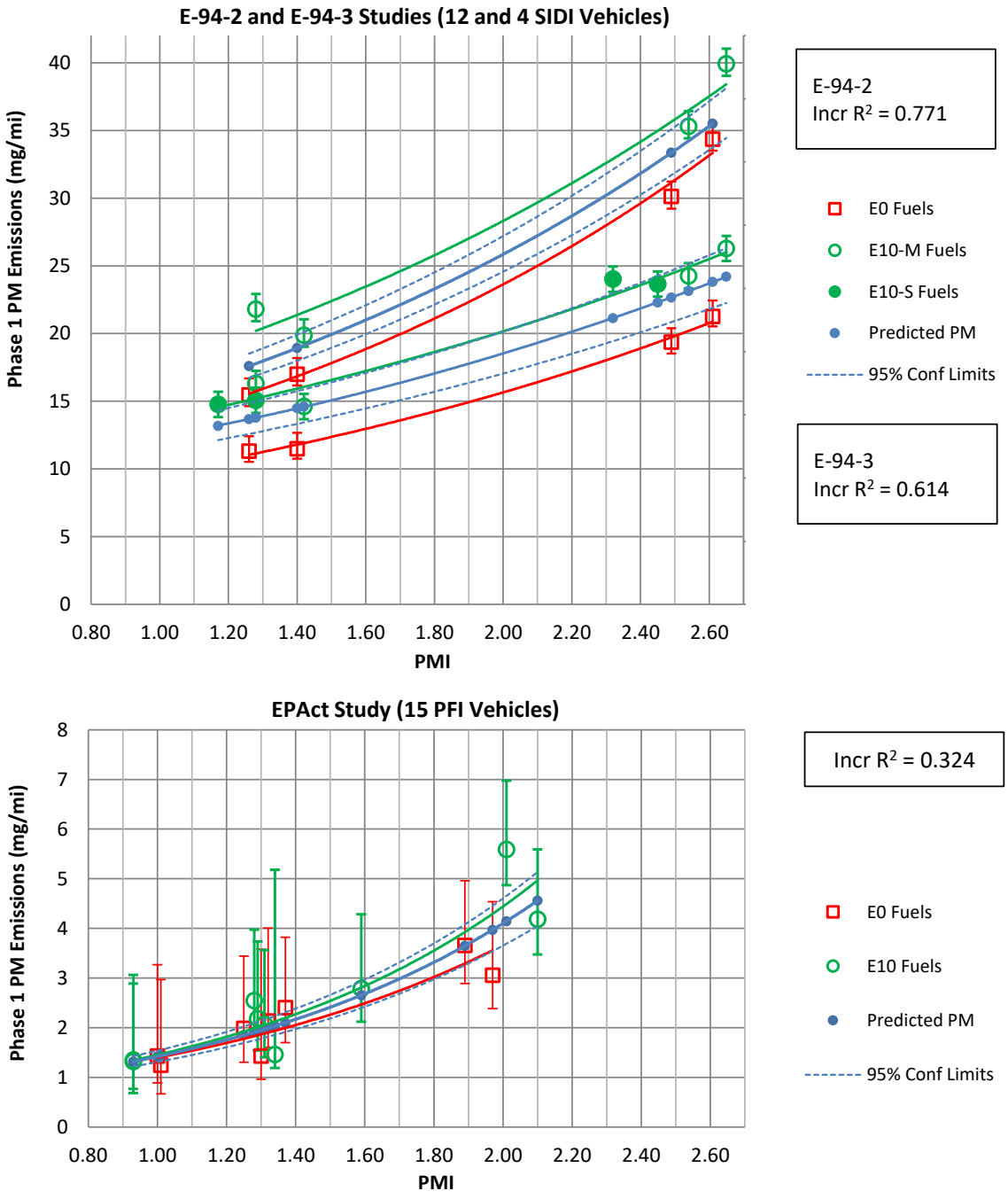
6.1 PMI Index

As shown throughout this section, the presence of ethanol in E10 fuels changes the relationship between the PM indices and Phase 1 PM emissions of the vehicles in ways that none of the indices can account for. This is illustrated clearly by its effect on the performance of the PMI.

Figure 6-1 presents the predictive performance graphically. For the E-94-2 and -3 studies (the top portion), the overall correlation lines "split the difference" between Phase 1 PM emissions of the E0 and the E10 fuels. All of the E10 fuels (green) lie above their correlation line (blue) while all of the E0 fuels (red) lie below the line. Separate trend lines computed using only the E10 and E0 fuels, and colored to match the fuels, parallel the overall correlation, but lie at or outside its 95 percent confidence limits. This fact, the directional consistency of the offsets for the fuels, and the sizes of the error bars on the individual data points all support the conclusion that Phase 1 PM emissions are different, and higher, for E10 fuels than for E0 fuels having the same value for the PMI. Formal tests for the statistical significance of the E10 versus E0 differences bear out this conclusion.

A similar offset is present in the correlation for the EPAct study fuels, but the emission differences from the correlation line are smaller and the error bars on the individual data points are now much larger than the E10 versus E0 offset itself. The perturbing effect of ethanol is much smaller and harder to detect at the much lower PM emission levels of the EPAct study's PFI vehicles.

Figure 6-1
Predictive Performance of PMI for E0 and E10 Gasolines



Overall, the E-94 data support the conclusion that ethanol at the E10 level increases Phase 1 PM emissions over and above the level that one would expect from the PMI value of a fuel. The nature of the offset appears to be one that could be represented by a percentage difference. The increased emissions for E10 fuels and the percentage nature of the offset are consistent with conclusions drawn in both the E-94-2 and E-94-3 studies.

Table 6-1 summarizes the predictive performance of the PMI for E0-E10 gasoline and can be compared to the prior Table 5-1 for E0 gasoline. Here, the table content is expanded to report the EtOH Bias affecting the emissions correlation. The EtOH Bias is calculated as the average predictive error (the residuals from the overall correlation line) made for the E10 fuels minus the average predictive error for the E0 fuels and, therefore, measures the size of the vertical gap between E10 and E0 fuels as a percent of the average PM emissions predicted by the correlation. It is reported for all fuels on average and for the subgroups of Low and High PMI fuels.[19]

Table 6-1
Summary of PMI Performance for E0-E10 Gasolines

	E-94-2	E-94-3	EPAct
Median Phase 1 PM Emissions (mg/mi)	25.3	19.1	2.4
Incremental R^2	0.771	0.614	0.324
RMS Error	11%	14%	19%
mg/mi	2.9	2.7	0.4
MAE Error	23%	26%	39%
mg/mi	5.7	5.0	0.9
PM index of the fuel	1.28	1.40	1.34
EtOH Bias: Avg. Predictive Error (E>0 minus E0 fuels)	19%	25%	10%
For Low PMI Fuels	24%	29%	6%
For High PMI Fuels	11%	21%	24%
Fuel variables adding explanatory power ($p \leq 0.01$)			
When EtOH term is required	Benzene	None	None

For the E-94-2 and -3 studies, the overall fit of PMI to the E0-E10 data is reduced compared to the fit for E0 fuels as measured by all performance metrics. Its Incremental R^2 values are down by 0.08 and 0.17 for E-94-2 and -3, respectively, its RMS Error is up to 2.7–2.9 mg/mi compared to 0.7–0.8 mg/mi for the E0 fuels, and its MAE is above 5 mg/mi compared to 1.1 mg/mi or less for the E0 fuel.

The values tabulated for the EtOH Bias give a good indication of the size of the systematic error in predicting Phase 1 PM emissions for the E10 versus E0 fuels. Considering all of

[19] The fuels are classified as Low or High PMI fuels based on the PMI, with fuels having PMI ≤ 1.80 being termed Low PMI fuels and those having PMI > 1.80 as High PMI fuels. This classification of fuels based on the PMI is used in the assessment of all indices.

the fuels, PM emissions from E10 fuels are 19% and 25% higher in the E-94-2 and -3 studies, respectively, than for the comparable E0 fuels. The difference attributed to the presence of ethanol in E10 fuels is substantial when it is remembered that, for the E0 fuels alone, the PMI index had RMS Errors of only ±2% and MAEs of only ±3%.

For the EPAct study, the index performance for E0-E10 fuels can be judged as equivalent to (or somewhat better than) that for the E0 fuels. The Incremental R^2 improves from 0.297 to 0.324, the RMS Error improves from 0.5 to 0.4 mg/mi, while the MAE is up only slightly from 0.8 to 0.9 mg/mi. The trend in the E10 versus E0 offset is also different than for the E-94-2 and -3 studies—it is small at low index values, but increases significantly as the index value increases to create an "expanding wedge".

It is difficult to judge whether the different shape ("expanding wedge") for the ethanol effect compared to the two E-94 studies is related to the PFI technology of the EPAct vehicles or is simply the result of random variability in a data set where there is limited ability to resolve the effect of ethanol at the low PM emission levels. The available data are not sufficient to test whether this is a technology difference (PFI versus SIDI). The second explanation (random variability where resolution is limited) is consistent with the much lower emission levels of the EPAct vehicles, a smaller mg/mi impact of ethanol, and a relatively larger role played by the random error in PM emissions measurements.

As a last step in the analysis, statistical tests were performed to identify fuel variables that might improve the performance of the indices. Given the important effect of ethanol, the tests were performed by requiring that ethanol be present as an explanatory variable; all remaining fuel property variables were then tested for statistical significance. When done for the PMI, the tests identified that Benzene could improve predictive performance for the E-94-2 study, but that no other variable could do so for the E-94-2 and EPAct studies. Given the absence of confirmation from two of the three studies, the Benzene identification for E-94-2 is unlikely to be meaningful; it could be a result of random chance given the large number of such statistical tests being made.

The chief conclusion drawn from these results is that ethanol can have a substantial effect on the ability of the PMI to predict Phase 1 PM emissions from a mixed group of E0 and E10 fuels. When applied to the E0 fuels (prior section), the index was able to accurately predict emissions for the E-94-2 fuels. Although the analysis did not consider a group of E10 fuels alone, the same should be true if it were applied to such a group, although the empirical coefficient between the index and emissions would be different. In fact, the vertical offset between the E10 and E0 fuels and the parallelism of the separate E10 and E0 trend lines for the E-94 studies in Figure 6-1 suggest that Phase 1 PM emissions in the E-94 studies may be proportional to the PMI value <u>but with different emission coefficients for the E10 and E0 fuels</u>.

The implication of this result is that the PMI may be subject to a substantial systematic error when applied to a mixed group of fuels. For example, it could place the fuels in an incorrect order if used to rank a group of E0 and E10 fuels for emissions. E10 fuels could be systematically ranked at lower emission levels and E0 fuels at higher emission levels

than would actually be observed. As shown, in the tables and figures of Appendix B, the same systematic error appears to affect all of the indices. Given this, the PMI will be more successfully used to guide the formulation of real world fuels when all share the same ethanol content.

6.2 Comparative Performance of the Indices

This section examines the comparative performance of the indices for E0–E10 gasoline. As shown, index performance is adversely affected when E10 fuels are included because the Phase 1 PM emissions of E10 and E0 fuels generally lie on opposite sides of the overall correlation line. In spite of this, some indices continue to perform relatively well, while the performance for others is severely affected.

6.2.1 Incremental R^2 Performance Measure

Figure 6-2 shows the comparative performance of the PM indices for E0-E10 gasoline as indicated by the Incremental R^2 measure. As before, the indices are presented in groups for the three studies and the columns shaded with horizontal bars to indicate indices that are statistically different (poorer) than the one best performer at 95% confidence. For this fuel group, the statistical uncertainty of this metric can be evaluated for the E-94-3 study because, with 12 fuels and 4 vehicles, is sample size is judged to be sufficiently large. Because additional fuel analyses were not possible for the EPAct study, PMI-A, 1/SP and OESI*-DHA MW could not be evaluated.

For the E-94-2 and -3 studies, five of the PM indices demonstrate consistently good performance in this comparison:

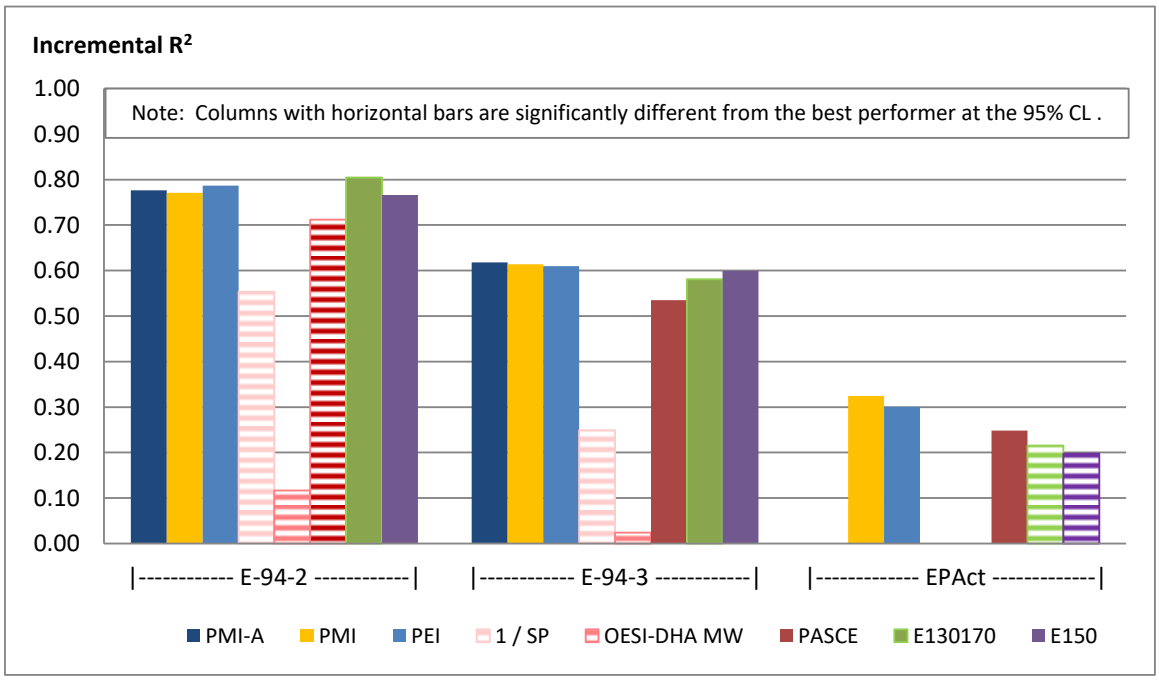
- PMI-A, PEI, and E130-170 are the high performers for the E-94-2 study, followed closely by PMI and E150.

- PMI-A, PMI, PEI, and E150 are the high performers for the E-94-3 study, followed closely by E130-170.

For E-94-2, the E130-170 index outperforms the others, followed by PEI and PMI-A as high performers, in that order. The Incremental R^2 values range from a low of 0.79 (PEI) to a high of 0.80 (E130-170). For E-94-3, the high performers in order are: PMI-A, PMI and PEI, and E150, with Incremental R^2 values that range from a low of 0.60 (E150) to a high of 0.62 (PMI-A). While E130-170 and E150 switch places as high performers for the two studies, the other index is not far behind.

PASCE trails these two indices by a modest margin and is just barely flagged as significantly different (poorer) that the one best performer. The other two indices, 1/SP and OESI*-DHA MW, are consistently flagged as statistically different (poorer) in this group where E0–E10 fuels are included.

Figure 6-2
Index Performance based on Incremental R^2: E0–E10 Gasoline



The levels of index performance for E0-E10 gasoline is below that for E0 gasoline in all cases as a direct result of the fact that E10 fuels have higher emissions than E0 fuels at any given index value. The high performers in the E-94-2 study achieved Incremental R^2 values of 0.84 to 0.85 for E0 gasoline, but only 0.77 to 0.80 for E0-E10 gasoline. For the E-94-3 study, the high performers achieved 0.78 to 0.79 for E0, but only 0.60 to 0.62 here for E0-E10 gasoline. The performance reductions are enough to be notable, but the levels seen here for the high performers still indicate good performance.

For the EPAct study, the PMI and PEI are the only high performers, while the other three indices lag behind; this also was seen for the E0 fuels. Here, the performance levels of the high performers are reduced when E10 fuels are included, but the indices based on hydrocarbon speciation continue to perform substantially better than correlation-based methods.

6.2.2 RMS Error Performance Measure

Figure 6-3 shows the comparative performance of the PM indices for E0-E10 gasoline as indicated by the RMS Error measure. As before, the indices are presented in groups for the three studies and the statistically significance differences are marked by horizontal shading. The pattern of high performers for this measure is generally similar to that for the Incremental R^2, but some differences exist, specifically:

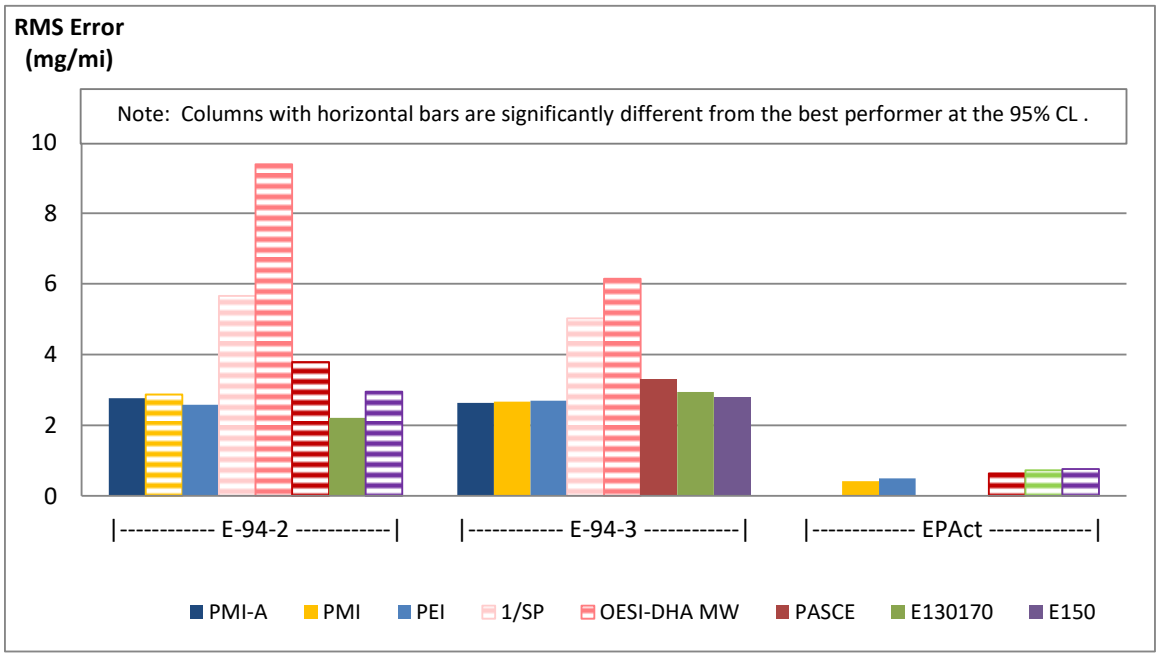
- For the E-94-2 study, E130-170 is the best performer with an RMS Error of 2.2 mg/mi, but it is followed closely by PEI and PMI-A. PMI is a close contender but

its RMS error is just high enough to be judged statistically different (poorer) than E130-170. These indices were also higher performers based on the Incremental R^2 measures.

- For the E-94-3 study, PMI-A, PMI, and PEI are the high performers, in order. They are followed very closely by E150, E130-170 and PASCE.

  For the EPAct study, the PMI and PEI are the high performers with RMS Errors of 0.5 mg/mi or less. The other three indices are statistically different (poorer) than the one best performer (PMI). One would expect the PMI-A to be among the high performers if it could be evaluated for the study.

Figure 6-3
Index Performance based on RMS Error: E0-E10 Gasoline



The PASCE index lags behind the high performers in two of the three groups. The 1/SP and the OESI*-DHA MW indices have the largest RMS Errors of any index for reasons that are shown in a later subsection.

Taken together, these results point to PMI-A, PMI and PEI—the indices based on hydrocarbon speciation—as being high performers on a consistent basis. The E130-170 is a close competitor and is the best performer the E-94-2 study. It is clear that a correlation-based method can perform well across a range of fuels.

6.2.3 MAE Performance Measure

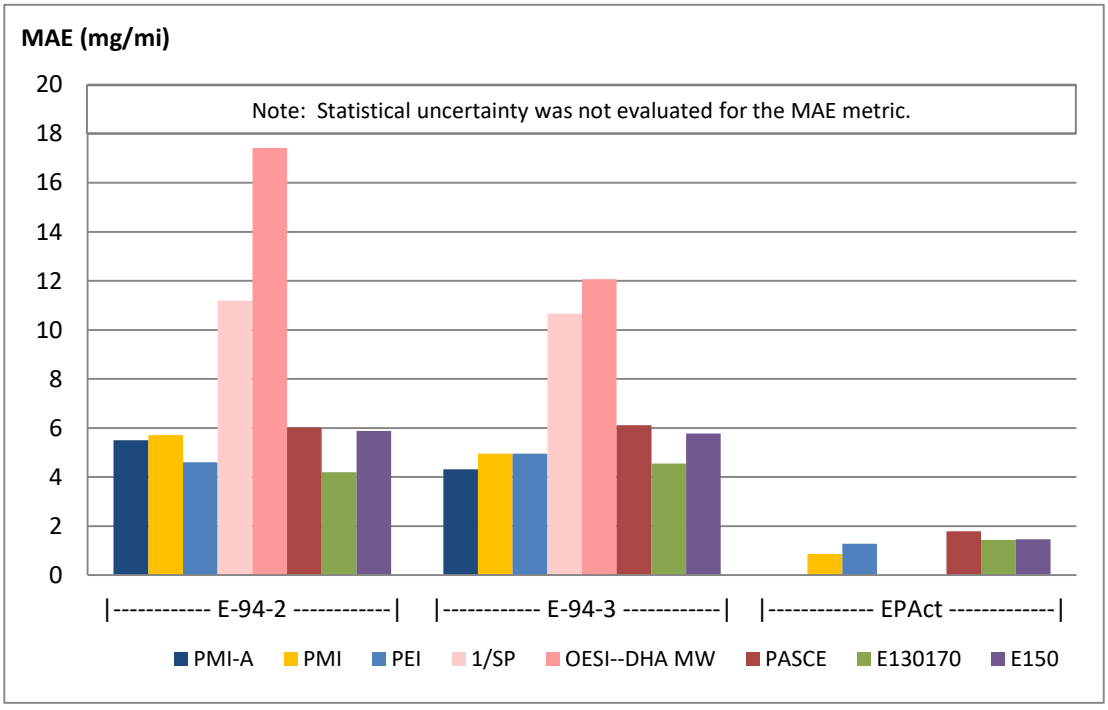
Figure 6-4 shows the comparative performance of the PM indices as indicated by the MAE measure. Here, the pattern of index performance is similar to that for the RMS Error measure. Specifically:

- For the E-94-2 study, E130-170 is the one performer with an MAE of 4.2 mg/mi, but it is followed closely by PEI and PMI-A among the top three. PMI is a close contender.

- For the E-94-3 study, PMI-A is the best performer with an MAE of 4.3 mg/mi, but it is followed closely by E130-170 as before and by PMI and PEI in a tie.

- For the EPAct study, the PMI is the best performer followed closely by PEI, both with MAEs of 0.5 mg/mi or less; both were the high performers for the RMS Error measure. One would also expect the PMI-A to be among the high performers if it could be evaluated for this study.

Figure 6-4
Index Performance based on MAE: E0-E10 Gasoline



6.3 Enhanced versus PMI

As in the assessment of E0 fuels, the PMI-A proves to perform as well as the PMI, but not clearly better for the group of E0-E10 fuels. Its performance for the E0 fuels in the group

is unchanged from that found earlier. Its performance for E10 gasoline is subject to the same perturbing influence of ethanol and to a comparable extent as PMI. This is not surprising as the difference between the two indices is related to the speciation of the heaviest hydrocarbons by the AVFL-29 DHA procedures and not to the treatment of ethanol in the calculation.

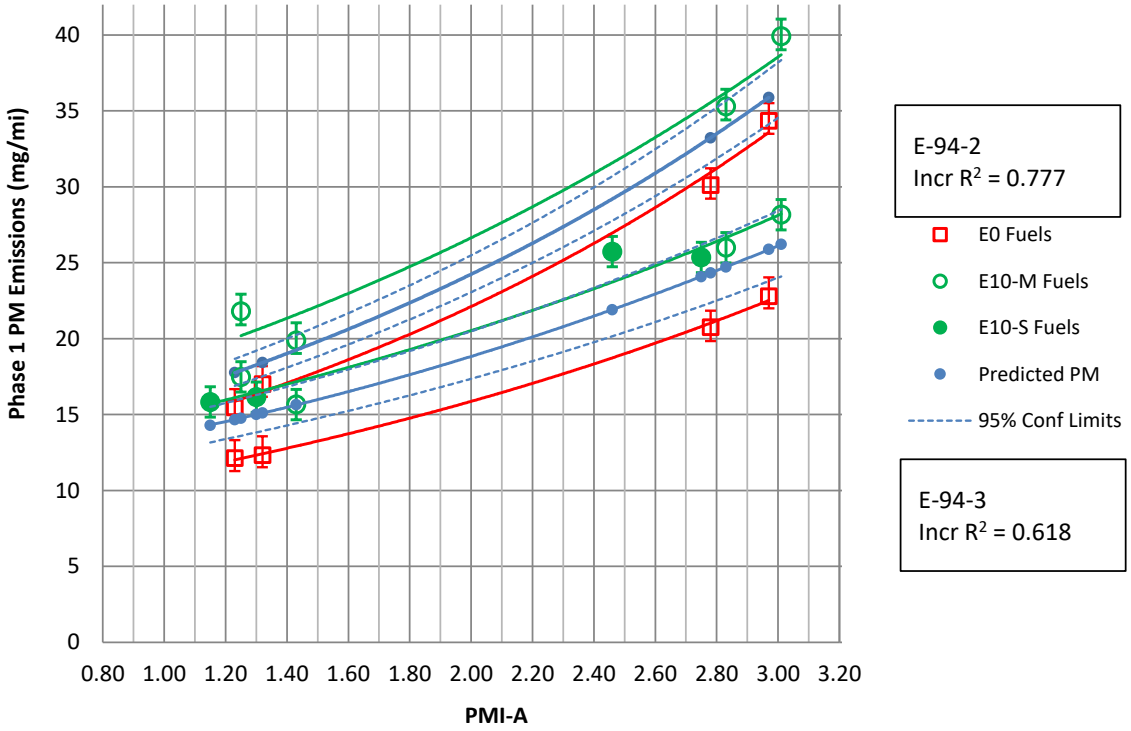
Figure 6-5 demonstrates the similarity of the predictive performance and the EtOH Bias the PMI-A (top portion) and PMI (bottom portion) using the E-94-2 and -3 studies for illustration. Except for the extension of the horizontal axis to higher values for the PMI-A, there is no appreciable difference in the quality of the correlations or the relationship of E0 and E10 fuels to the correlations. Table 6-2 provides further support; as easily seen, there is no material difference in the performance of the two indices across the performance measures.

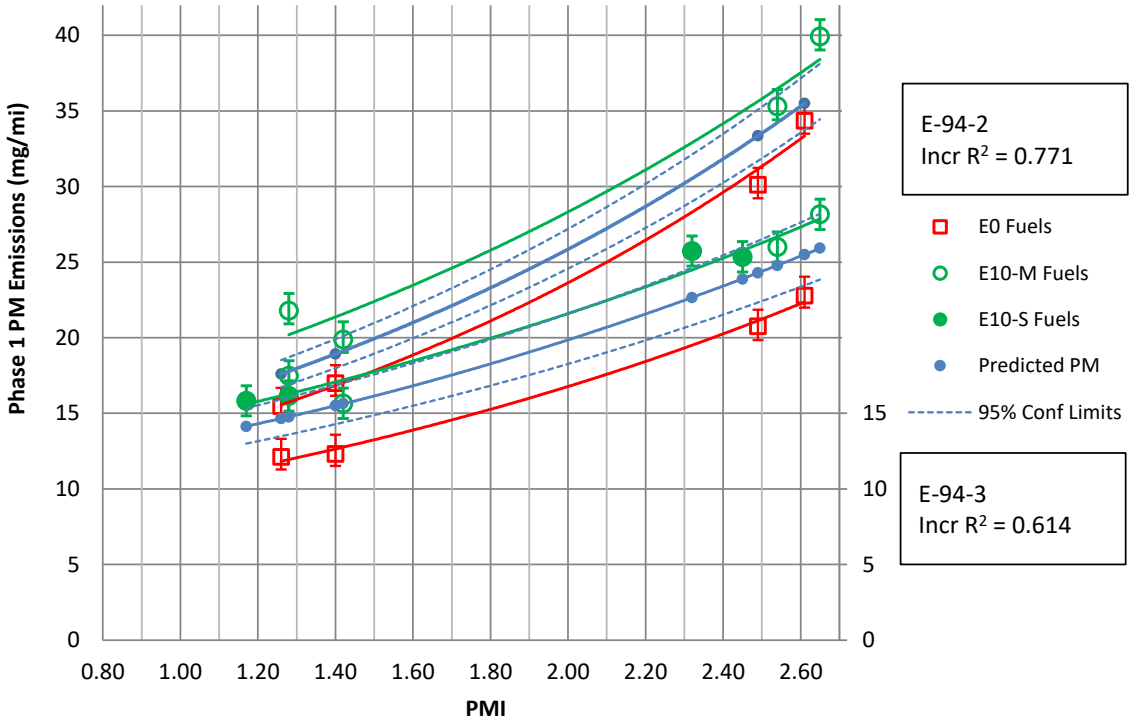
Table 6-2
Performance Comparison for PMI and PMI-A: E0-E10 Gasoline

	E-94-2 Study		**E-94-3 Study**	
	Original PMI	PMI-A	PMI	PMI-A
Incremental R^2	0.771	0.777	0.614	0.618
RMS Error	11%	11%	14%	14%
mg/mi	2.9	2.8	2.7	2.6
MAE Error	23%	22%	26%	23%
mg/mi	5.7	5.5	5.0	4.3
PM index of the fuel	1.28	1.25	1.40	1.32
EtOH Bias: Avg. Predictive Error				
(E>0 minus E0 fuels)	19%	18%	25%	24%
For Low PMI Fuels	24%	22%	29%	27%
For High PMI Fuels	11%	11%	21%	22%

As has been noted, PMI and PMI-A are highly correlated variables for the E-94-2 and -3 fuels. In circumstances where two variables are highly correlated, one variable (PMI-A) can be a more complete and reliable assessment of a characteristic, but it may not demonstrate notably better predictive performance than another variable (PMI). For typical gasolines, blended from refinery hydrocarbon streams with attention to commercial specifications but not to the content of the heavy versus the heaviest hydrocarbons, the two indices can have comparable ability to predict PM emissions. In other cases, such as fuels where the content of the heaviest aromatic hydrocarbons is atypically higher (or lower), the PMI-A could be a clearly better and more reliable predictor of emissions.

Figure 6-5
Predictive Performance of the PMI and PMI-A Indices for E0-E10 Gasoline
E-94-2 and E-94-3 Studies





This is what is seen in the performance data, although only to a small extent. PMI and PMI-A show essentially the same performance based on the Incremental R^2 and RMS Error, but the PMI-A shows slightly better performance based on MAE (the worst-case error) as one might expect from its more comprehensive speciation of the fuel hydrocarbons.

6.4 The Nature of the Ethanol Bias

The inclusion of E10 fuels causes the performance of all indices to decline by varying amounts. Ethanol will have an effect on the value computed for most of the indices, whether through its inclusion in the speciation of hydrocarbons or its effect on bulk fuel properties used in the indices. The PEI index is affected by the reduced prevalence of the aromatic hydrocarbon groups in the fuel (even though ethanol is not included in the formula). While all of the indices exhibit an ethanol bias, particularly in the E-94 data, its nature can differ among the indices. This section examines information pertaining to the two E-94 studies to demonstrate this. The indices are grouped to simplify the presentation.

6.4.1 PMI, PMI-A, and PEI

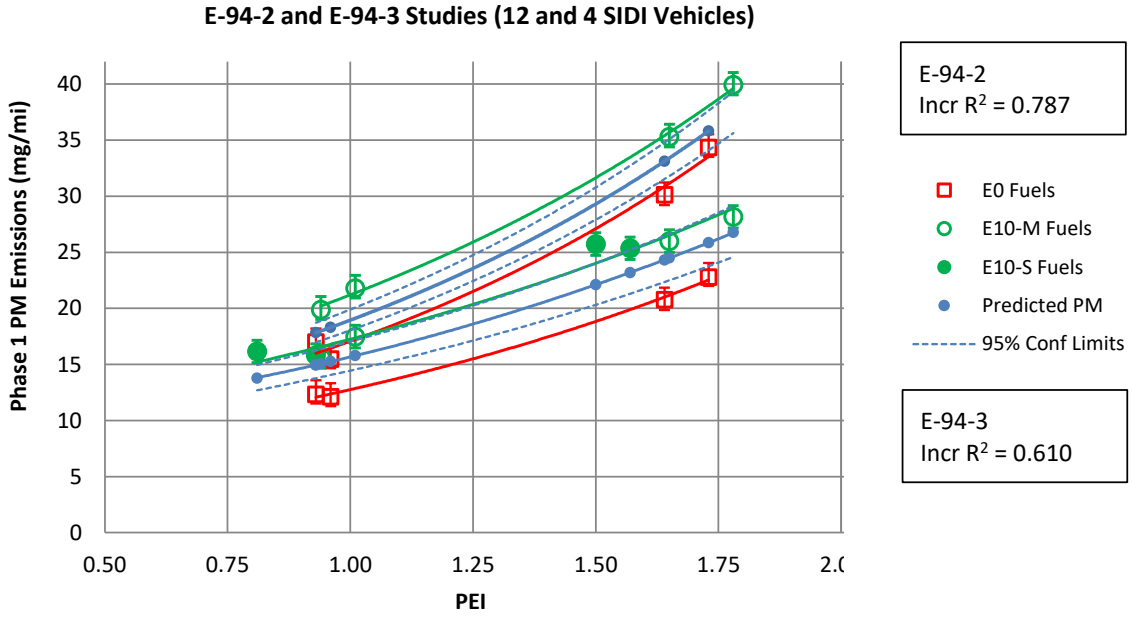
The PMI, PMI-A, and PEI are the three indices that are based on a hydrocarbon speciation. The previous Figure 6-5 demonstrated that the ethanol bias in PMI and PMI-A is virtually indistinguishable and has the character of a vertical offset in PM emissions (E10 emissions > E0 emissions). As Figure 6-6 shows, PEI (whose performance closely matched that of the two PMIs) has a similar ethanol bias in the E-94 studies. The vertical offset is consistent across the range of index values and could be presented by a percentage difference.

6.4.2 Ethanol Bias for 1/SP and OESI*-DHA MW

These related indices prove to have the lowest performance level for the E0-E10 fuels among the eight indices examined. Figure 6-7 shows their predictive performance graphs for the two E-94 studies. The 1/SP index is given in the top portion and the OESI*-DHA MW index in the lower portion. Here, the two E-94 studies have been plotted on separate axes with the axis for E-94-2 on the left and the axis for E-94-3 on the right so that both studies can be shown on one plot without overlap.

This complicated graph demonstrates the complex relationship between Phase 1 PM emissions and the two indices for the E0 and E10 fuels. For 1/SP, the offset between the E0 and E10 trend lines is more than just a percentage difference for the E0 and match-blended E10 fuels in the E-94-2 study. In fact, different intercepts and slopes would be required to give the best representation of PM emissions for the E0 and E10 fuels. Essentially, separate correlations are required for the two groups. On the other hand, a percentage offset appears to be a suitable representation for the E-94-3 study with its mix of E0 and match- and splash-blended E10 fuels. In both cases, individual fuels fall far from their trend lines in many cases. This was also seen for the E0 fuels (see Figure 5-5) but the scatter here is more pronounced.

Figure 6-6
Ethanol Bias of PEI for E0-E10 Fuels



The OESI*-DHA MW index is the product of the 1/SP index and terms that represent the molecular weight of the fuel and its stoichiometry based on the carbon, hydrogen and oxygen content. For this index, different intercepts and slopes would be required in both E-94 studies to give the best representation of emissions for the E0 and E10 fuels. The slopes for E0 and E10 fuels are very different for the E-94-2 fuels, but nearly parallel for the E-94-3 fuels. Although smoke point was measured for E0 and E10 fuels using the same laboratory procedure, its relationship to PM emissions appears to be substantially different for the E0 and E10 fuels as groups. The additional terms included in OESI*-DHA MW appear to increase the differences between the relationships for E0 and E10 fuels.

6.4.3 Ethanol Bias in the PASCE Index

The PASCE index is a correlation-based method that uses the C/H Ratio and the E170 distillation point as predictors. Figure 6-8 shows the predictive performance graph for PASCE for the two E-94-2 studies. Its performance for the E0-E10 fuels is good, but it lags behind the leading indices because it has more scatter. The nature of the ethanol bias is the same as for the three indices based on hydrocarbon speciation. It could be represented by a constant offset.

Figure 6-7
Ethanol Bias of 1/SP and OESI*-DHA MW for E0-E10 Fuels
(E-94-2 and E-94-3 Studies)

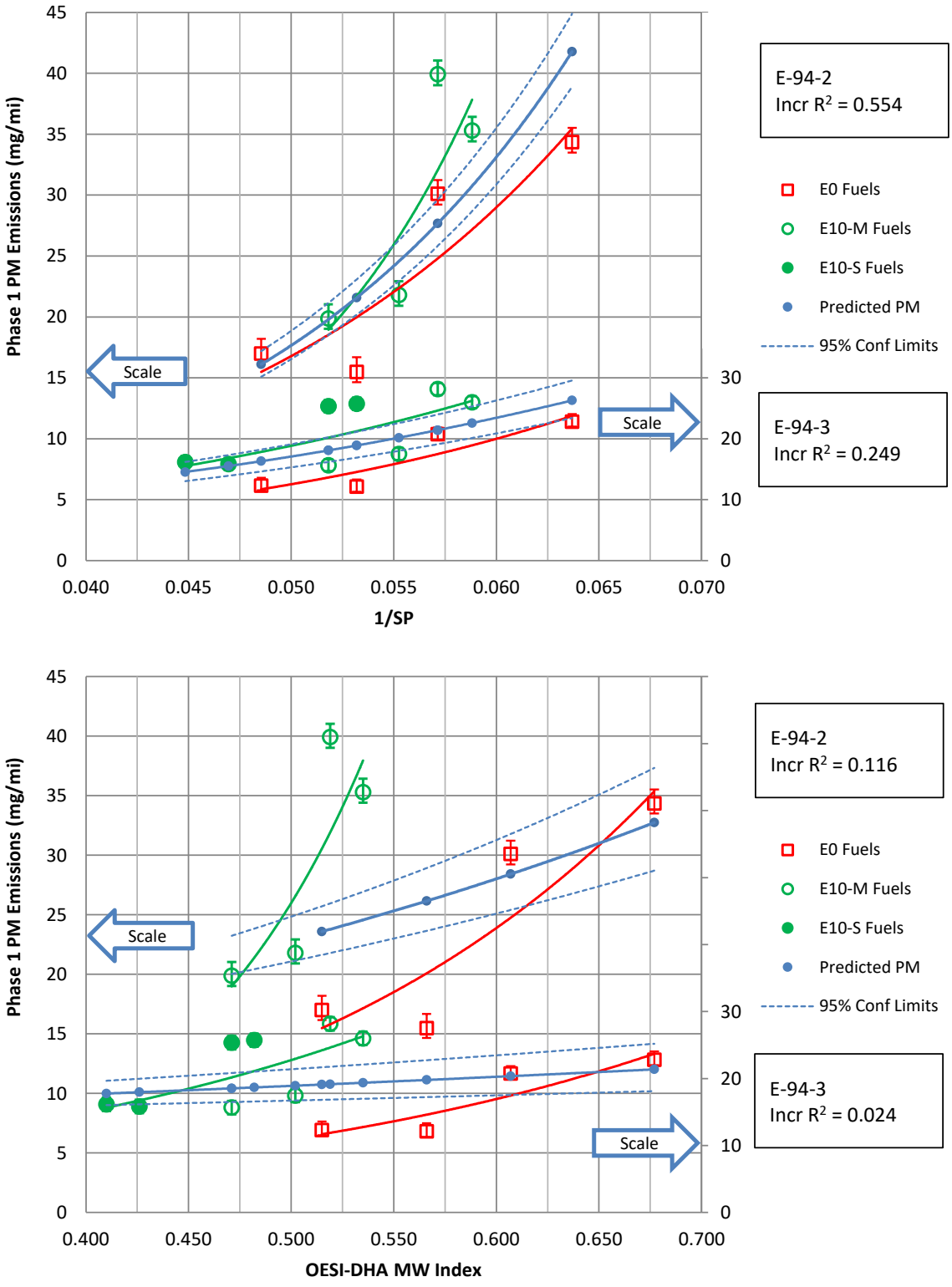
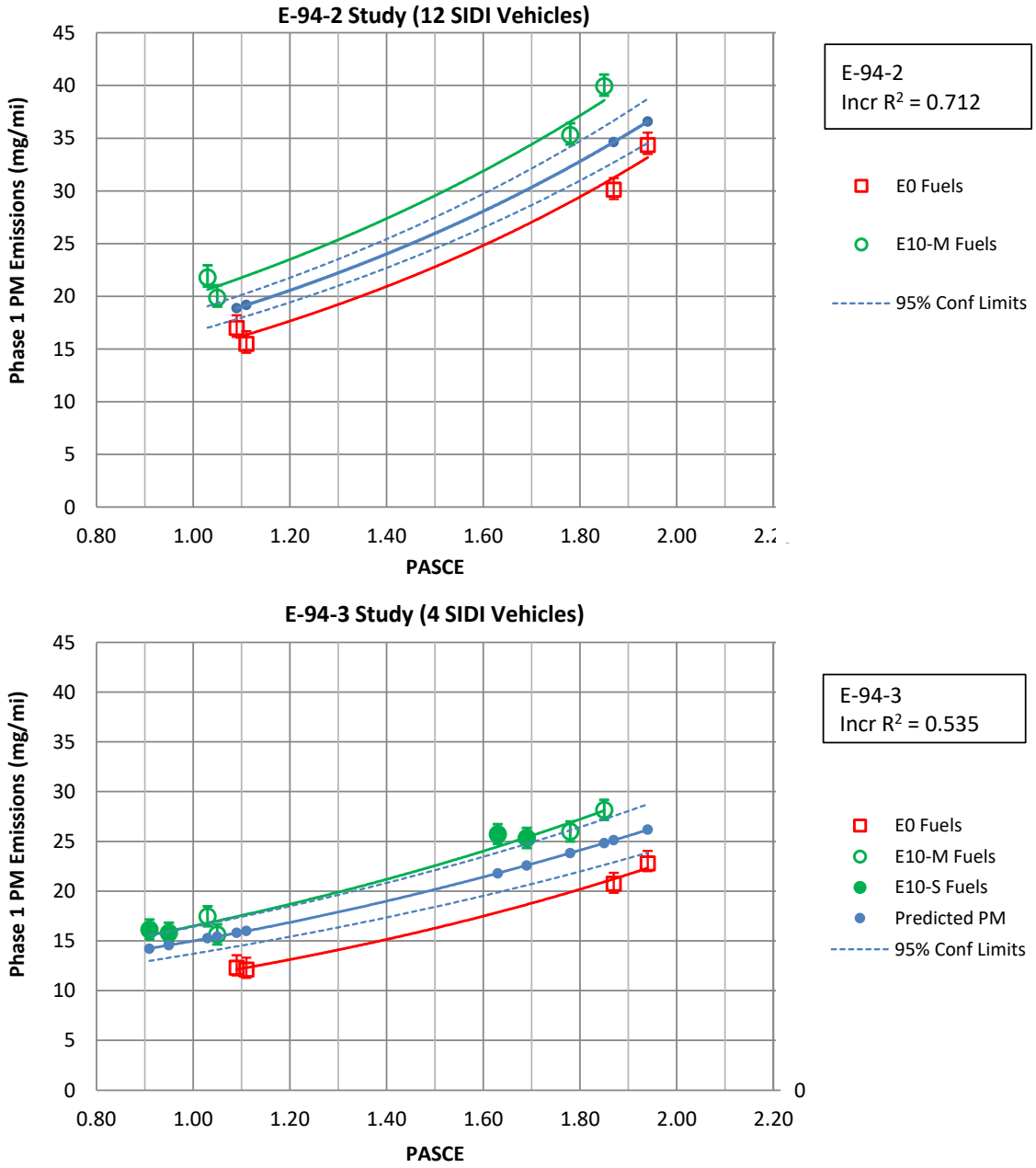


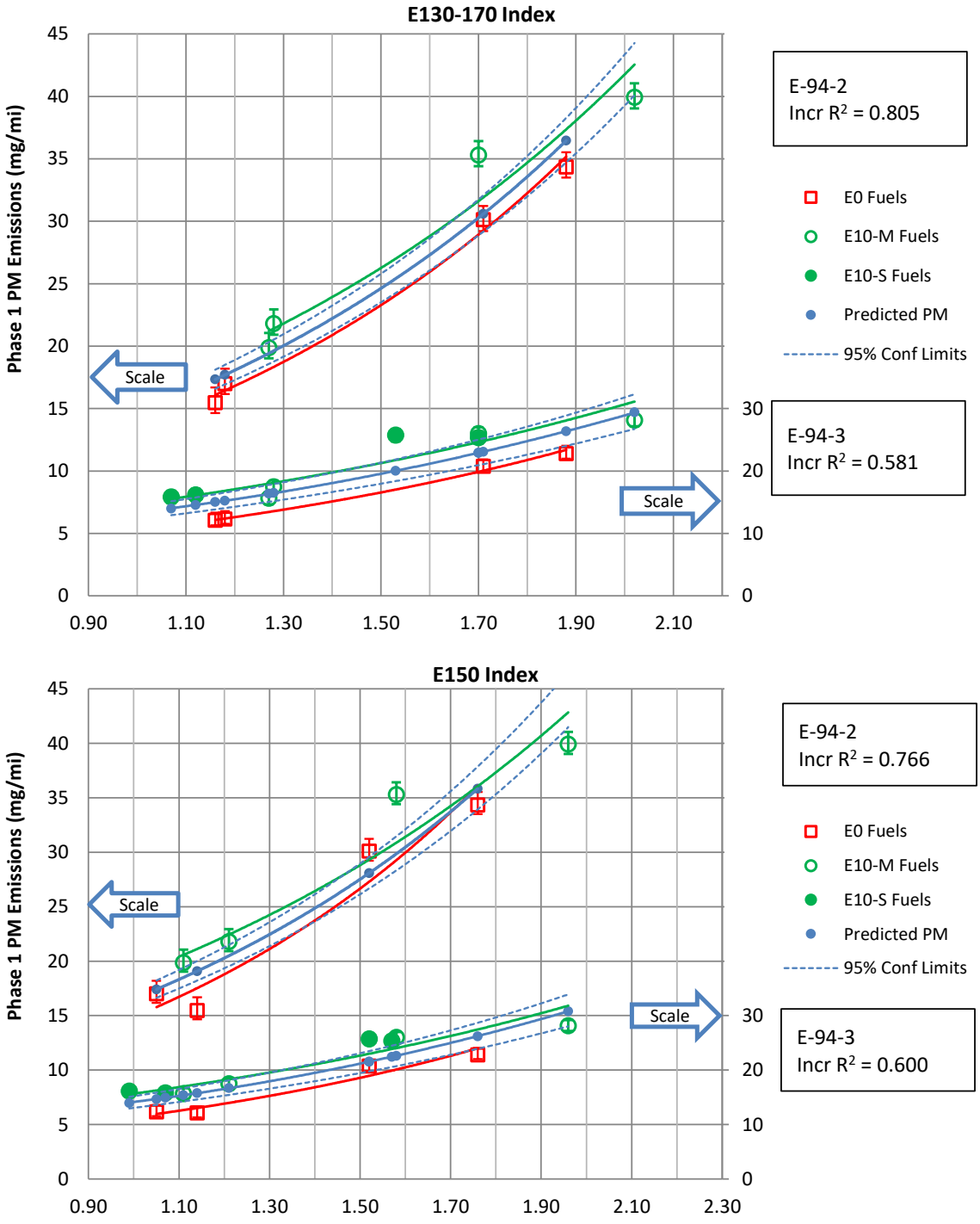
Figure 6-8
Ethanol Bias of PASCE for E0-E10 Fuels



6.4.4 Ethanol Bias in the E130-170 and E150 Indices

The E130-170 and E150 indices are correlation-based methods that use only selected points on the distillation curve. Figure 6-9 shows the predictive performance graphs for these indices for the two E-94 studies. The E130-170 index displays a similar vertical offset in

Figure 6-9
Ethanol Bias of the E130-170 and E150 Indices
(E-94-2 and E-94-3 Studies)



Phase 1 PM emissions between the E0 and E10 fuels. This is seen in both studies and could be represented by a percentage difference.

The result is not as clear cut for the E150 index, which is the simplest index examined in this study. For the E-94-2 study, the index displays an offset between emissions for E0 and E10 fuels that is larger at low index values and converges at high index values. For the E-94-3 study, the emission trend lines for E0 and E10 fuels are more nearly parallel, but still show some narrowing at high values. There is also more scatter between the trend lines and the individual E0 and E10 data points. The ethanol bias of this index might or might not be adequately represented by a percentage difference. Additional analysis would be required to judge whether the apparent convergence is meaningful (statistically significant) or could be the result of random chance.

6.4.5 Summary of Ethanol Bias in the PM Indices

All of the PM indices are affected by an ethanol bias in which Phase 1 PM emissions of the E10 fuels are found to be consistently higher at any given index value than the emissions of E0 fuels. If used to rank fuels based on expected Phase 1 PM emissions, the indices would generally rank E10 lower (better) than they should be and the E0 fuels higher (worse) than they should be. Table 6-3 summarizes the characteristics of the ethanol bias for the eight indices based on the two E-94 studies. For most of them, the ethanol bias takes the form of vertical offsets with a size of 10 percent or more.

Table 6-3
Characteristics of the Ethanol Bias in PM Indices

PM Index	Average Ethanol Bias (%)		Trend of Ethanol Bias
	E-94-2	E-94-3	
PMI	19	25	Vertical Offset
PMI-A	18	25	Vertical Offset
PEI	18	26	Vertical Offset
1/SP	38	30	Complex Relationship
OESI*-DHA MW	20	32	Complex Relationship
PASCE	26	37	Vertical Offset
E130-170	11	24	Vertical Offset
E150	10	23	Converging Offset

All of the indices show a smaller bias for the E-94-2 fuels and a larger one for the E-94-3 fuels. The E-94-2 fuels include four E0 and four match-blended E10 fuels; the E-94-3 fuels include the same fuels plus four additional splash-blended E10 fuels. While the match-blended E10 fuels were controlled to meet the same commercial specifications for gasoline as the E0 fuels, the splash-blending performed in E-94-3 allowed fuel properties to vary freely in response to the addition of ethanol. As a result, the splash-blended E10

fuels could depart to some extent from commercial specifications for AKI and vapor pressure.

PMI, PMI-A and PEI—the indices based on hydrocarbon speciation—share a similar ethanol bias. The bias takes the form of a vertical offset that could be represented by a percentage difference with similar sizes for the two studies: 18-19 percent for the E-94-2 study and 25-26 percent for the E-94-3 study. The PMI indices count ethanol as making a near-zero contribution to PM emissions because its high vapor pressure minimizes its "i-term" in the Honda Equation. The PEI index counts ethanol as not contributing because the index considers only the contributions from aromatic hydrocarbons. As a result, the effect of ethanol on the PM index values for ethanol-containing fuels is determined by the dilution of the gasoline hydrocarbons. It is not surprising that the three indices display similar ethanol biases.

The E130-170 and E150 indices are the two best performing indices with respect to the ethanol bias, particularly for the E-94-2 fuels meeting commercial gasoline specifications. For these indices, the bias could be presented by a vertical offset of 10-11 percent for the E-94-2 fuels and of 23-24 percent for the E-94-3 fuels. E130-170 frequently proved to be a high performer based on the three performance measures. Part of this performance and its lesser ethanol bias may be the fact that its distillation points were well-chosen to be diagnostic of the ethanol content of the fuels (E130) and the content of heavy hydrocarbons (E170). In fact, the E170 distillation point matches the 443K (170°C) temperature at which the vapor pressure is evaluated in the Honda Equation. Its predictive variables are directly affected by the ethanol content and the prevalence of heavy hydrocarbons in a fuel and, thus, may serve well as a two-variable surrogate for a detailed hydrocarbon speciation.

The other indices do not perform as well. The 1/SP and OESI*-DHA MW indices have a complex relationship between Phase 1 PM emissions and the index values. These relationships would require independent intercept and slope coefficients for each level of ethanol (E0, E10, etc.) to represent the PM emissions of the fuels.

The PASCE study has one of the largest ethanol biases. Its bias takes the form of a vertical offset equal to 26 percent on average for E-94-2 and 37 percent for E-94-3. PASCE uses the C/H ratio and the E150°C distillation point as predictor variables. The first is only weakly sensitive to the presence of ethanol in a fuel and the second is not as closely tuned to ethanol as the 130°C distillation point. Thus, its performance trails most of the other indices with respect to representing the PM emissions of both E0 and E10 fuels.

6.5 How the Indices Might be Improved

The last step in the analysis for each index was to search for additional fuel variables that might improve the performance of the indices for E0-E10 gasoline. This was done using stepwise regression techniques in which the residuals of the index correlations were tested against a list of available fuel properties. In this, an ethanol term was added to the base model to account for a vertical offset in emissions between E0 and E10 fuels. The term uses the ethanol content of the fuels as a continuous variable, but it acts much like a dummy

variable because the E0 fuels have zero ethanol content and the E10 fuels cluster around 9.5% ethanol content.

Table 6-4 summarizes the variables that were identified in this process. A large number of statistical comparisons are made in the process: about 20 for each PM index and emissions study for a total of about 450. Therefore, false positives are likely to occur in spite of the tightened $p \leq 0.01$ level that was required. Some of the variables indicated in the table are likely to be the result of random chance (a false positive).

In this assessment, Benzene was identified as a variable that might improve PMI and PMI-A for the E-94-2 fuels. Olefins was identified in two other cases (PASCE for the EPAct study and E130-170 for the E-94-2 study). Neither identification is likely to be meaningful as they are not confirmed in the other two studies and Benzene and olefins are gasoline components present in small quantities. While they could have emissions effects that are not accounted for by the indices, it is more likely they are either: (a) surrogates for selected fuels that demonstrated emission differences; or (b) false positives that occur merely by chance.

Table 6-4
Fuel Property Variables That Might Improve the Indices
Once Ethanol is Accounted for in Combined E0-E10 Evaluations

	E-94-2 Study	**E-94-3 Study**	**EPAct Study**
PMI-A	Benzene	—	(Not Evaluated)
PMI	Benzene	—	—
PEI	—	—	T60
1/SP	AKI, RVP, Aromatics, Density	< C10 Aromatics	(Not Evaluated)
OESI*-DHA MW	C10+ Aromatics	C10+ Aromatics	(Not Evaluated)
PASCE	—	—	DVPE, Olefins
E130-170	Olefins	—	Aromatics
E150	Aromatics, T60, Density	—	Aromatics

In other cases, the variable identifications may prove helpful in future work to improve the performance of the indices and the correlation-based indices in particular. Aromatics content, either in total (Aromatics) or subdivided into lighter (<C10 Aromatics) and heavier (C10+ Aromatics) components, are the variables most commonly identified. Density, a related variable, is identified in two other cases. These variables provide additional

information on the prevalence of aromatics and light versus heavy compounds for indices that do not otherwise have compositional information.

As for the E0 fuels, the indices could generally benefit from the addition of one or more variables in representing PM emissions of the EPAct fuels. The variables identified include T60, DVPE, Aromatics, and Olefins. The greater number of variables is thought to be the result of the design and blending of the EPAct fuels.

6.6 Summary of Index Performance for E0–E10 Gasoline

Table 6-5 compiles a summary of index performance for the E0-E10 gasolines in an effort to find consensus. The high performing indices for each measure are indicated in the table by their rank (including ties).

The three indices based on a hydrocarbon speciation—PMI-A, PMI and PEI—are again the ones most often classified as high performers, although the E130-170 index is the best performer for the E-94-2 study where its lower ethanol bias lends it an advantage. As before, PMI and PEI are the only indices classified as high performers for the EPAct study; we would expect PMI-A to join them if it could be evaluated for the EPAct fuels. Among the high performers, PEI receives the most high performer classifications (9), followed by PMI (7), PMI-A (6), and by E130-170 (3).

While the three speciation-based indices—PMI and PMI-A and PEI—and the E130-170 index are consistent high performers for E0-E10 gasolines, neither they nor the other indices are unbiased predictors of Phase 1 PM emissions for E0 versus E10 gasolines. The presence of a substantial ethanol bias in these (and other) indices means that further index development is needed before the indices can be used in real-world settings that involve emissions trade-offs between E0 and E10 gasolines.

Table 6-5
High performing Indices for E0-E10 Gasoline

		R^2	RMS Error	MAE
PMI-A	E-94-2	#3	#3	#3
	E-94-3	#1	#1	#1
	EPAct	(Not Evaluated)		
PMI	E-94-2			#2
	E-94-3	#2	#2	#3
	EPAct	#1	#1	#1
PEI	E-94-2	#2	#2	#2
	E-94-3	#2	#2	#3
	EPAct	#2	#2	#2
1/SP	E-94-2			
	E-94-3			
	EPAct	(Not Evaluated)		
OESI*—DHA MW	E-94-2			
	E-94-3			
	EPAct	(Not Evaluated)		
PASCE	E-94-2			
	E-94-3			
	EPAct			
E130-170	E-94-2	#1	#1	#1
	E-94-3			#2
	EPAct			
E150	E-94-2			
	E-94-3	#3	#3	
	EPAct			

7. PREDICTIVE PERFORMANCE FOR EPACT GASOLINES (E0-E20)

This section examines the performance of the indices for the group of E0-E20 gasolines in the EPAct study, which is the only one of the three emission studies to include E15 and E20 fuels. In comparison to the E-94 studies, the EPAct dataset involves PFI vehicles with much lower Phase 1 PM emission levels and fuels that were formulated with different objectives. Thus, the study offers a broader view of the range of ethanol content but for a different set of vehicles and fuels.

The lower emission levels of the PFI vehicles lead to a relatively greater role for the random errors of measurement that contribute to large error bars around the measured emissions. As a result, it is more difficult to resolve the effects of fuels and ethanol content on emissions. For this reason, only an abbreviated presentation of comparative performance results is given that focuses on the overall performance of the indices for E0-E20 gasoline in comparison to their performance for the group of E0-E10 gasoline. Appendix C contains complete results for the indices.

7.1 Comparative Performance of the Indices

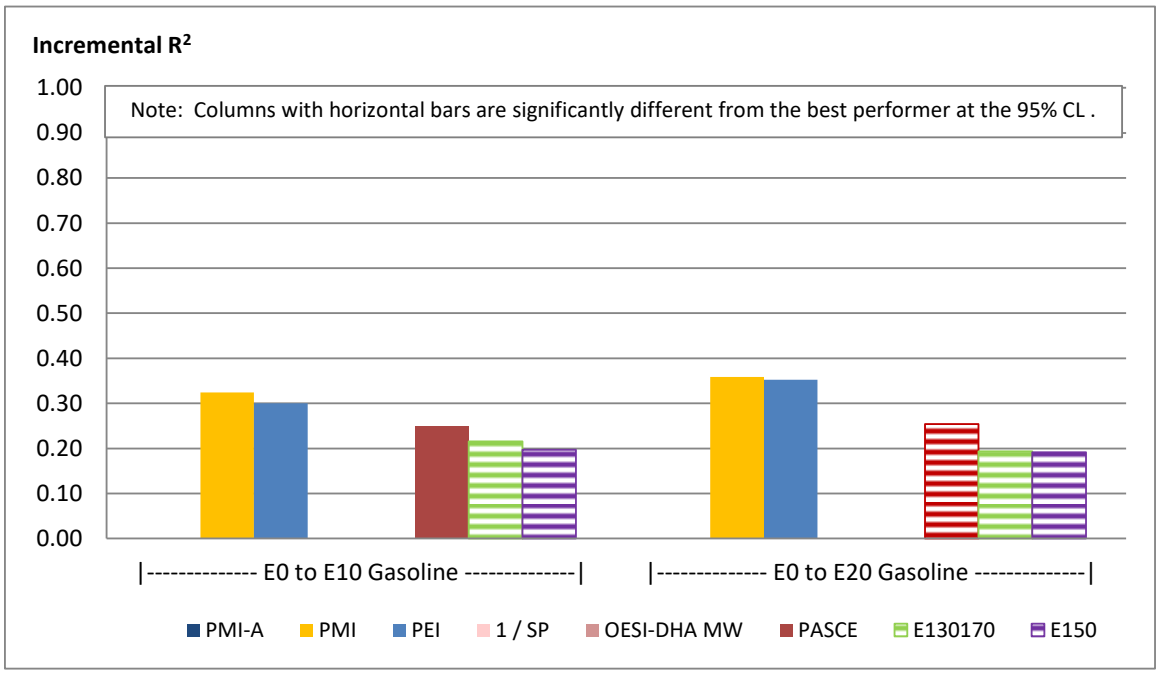
This section examines the comparative performance of the indices for E0-E20 fuels based on the three primary performance metrics. As shown, extending the range of ethanol content in the fuels leads to a similar outcome as previously seen for the E0-E10 fuels in the EPAct study and extends the findings to the E20 level.

7.1.1 Incremental R^2 Performance Measure

Figure 7-1 shows the comparative performance of the PM indices for E0-E20 gasoline based on the Incremental R^2 measure. The indices are presented in two groups with the group of E0-E10 gasolines on the left and the E0-E20 gasolines on the right. PMI-A and two related indices (1/SP and OESI*-DHA MW) could not be evaluated because additional fuel analyses were not possible.

The PMI and PEI indices are the two best performing indices for both groups of fuels—the E0-E10 gasolines and E0-E20 gasolines—with a small, but clear-cut margin over the other, property-based indices (PASCE, E130-170, and E150). PASCE is statistically different (poorer) than the one best performer for the E0-E0 group, while E130-170 and E150 are statistically different (poorer) in both groups. This difference between speciation- and property-based indices was seen in earlier results and likely due to the complex design of the EPAct fuels. One would expect PMI-A, the other speciation-based index, to rank with the best performers if it had been possible to evaluate it.

Figure 7-1
Index Performance based on Incremental R^2: EPAct Fuels through E20

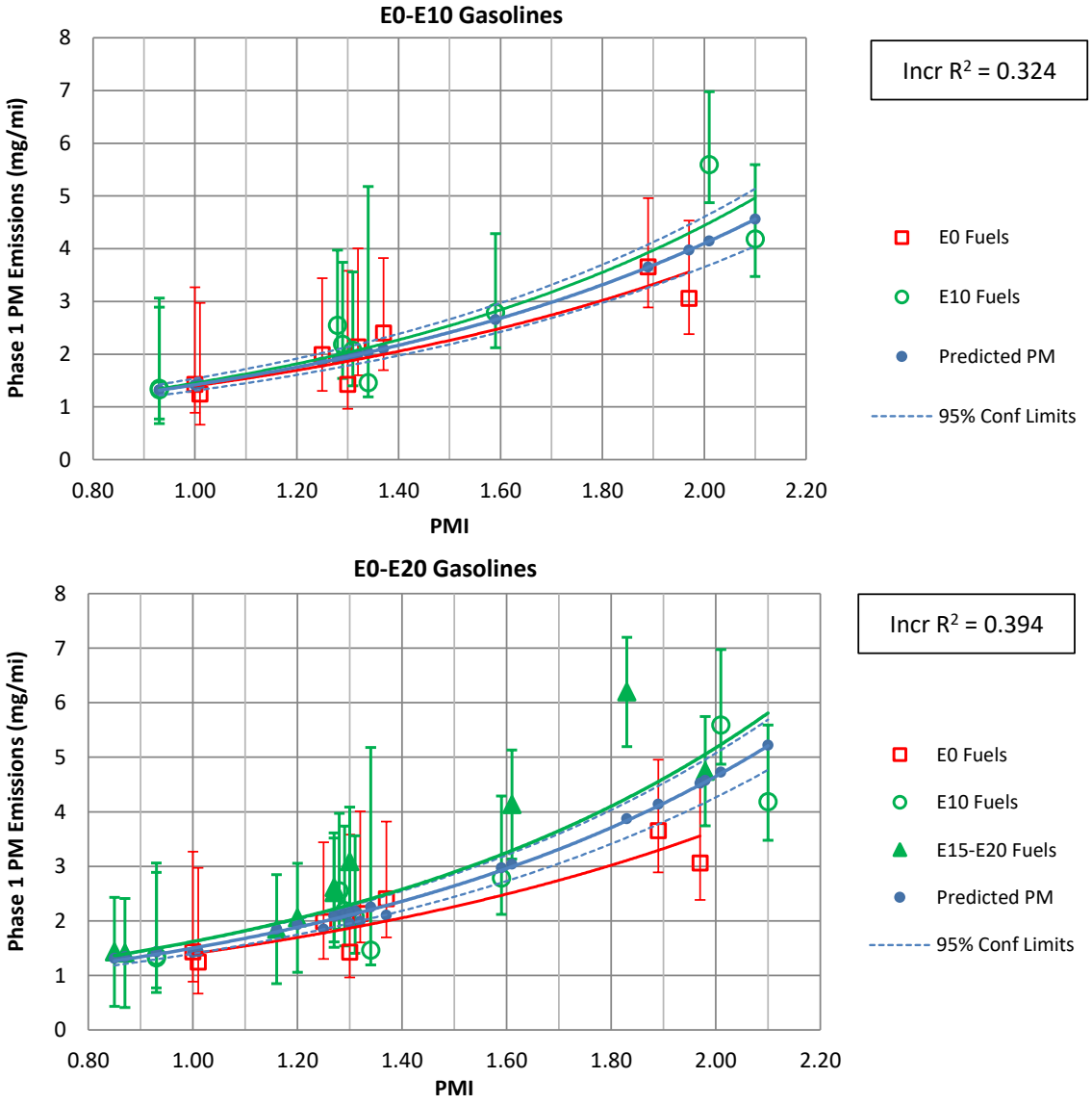


The other notable trend for the Incremental R^2 performance measure is that all of the indices perform better for the E0-E20 fuels than for the E0-E10 fuels. As shown, this has less to do with better performance of the indices for E15 and E20 gasolines than with differences in the two datasets. As a result, all of the indices benefit and their Incremental R^2 values improve.

Figure 7-2 presents the performance graphs for the PMI for the two sets of EPAct fuels. For E0-E10 gasolines (the top portion), all but one of the 17 data points overlaps the correlation line (in blue) in that the one-sigma error bars cross the line. For these 16, the measured emissions do not differ from the emissions predicted by the correlation line in either a statistical and practical sense. The error in prediction is counted for all data points in determining the Incremental R^2 value, but the one E0 fuel at PMI = 2.01 that lies well above the correlation line will make the largest contribution.

For E0-E20 gasolines (the bottom portion), we see a generally similar pattern. Here, all but one of the 27 data points overlaps the correlation line. The inclusion of E>10 fuels has steepened the slope of the line enough that the error bars of the E0 point at PMI = 2.01 now reach to the line. The one outlying data point is the E20 gasoline at PMI = 1.83. It now makes the largest contribution in determining the Incremental R^2 value but, here, it is only 1 of 27 data points. As a result, the calculated measure of the quality of fit improves, although inspection of the performance graphs reveals relatively little to prefer in the outcome for the E0-E20 gasolines versus the outcome for the E0-E10 gasolines.

Figure 7-2
Performance Graphs for PMI: EPAct Fuels



Lastly, the performance graphs reveal that the ethanol bias of this index for the EPAct fuels appears to take the form of an "expanding wedge" in which the PM emissions difference between E>0 and E0 fuels essentially vanishes at low PMI values but expands in size as the PMI increases. On average, the E>0 fuels lie above the correlation line while the E0 fuels tend to lie below, but there is considerable scatter in the data due to the relatively large effect of measurement errors at these low PM emissions levels. The vertical width of the "wedge" is greater for the E0-E20 fuels, which is consistent with the hypothesis that ethanol increases PM emissions (at any given PMI level) in proportion to its prevalence in E10 to E20 fuels. In spite of the considerable scatter in the data, tests indicate that EtOH is a statistically significant explanatory variable for the observed emissions differences, consistent with the emission difference being higher when the ethanol content is greater,

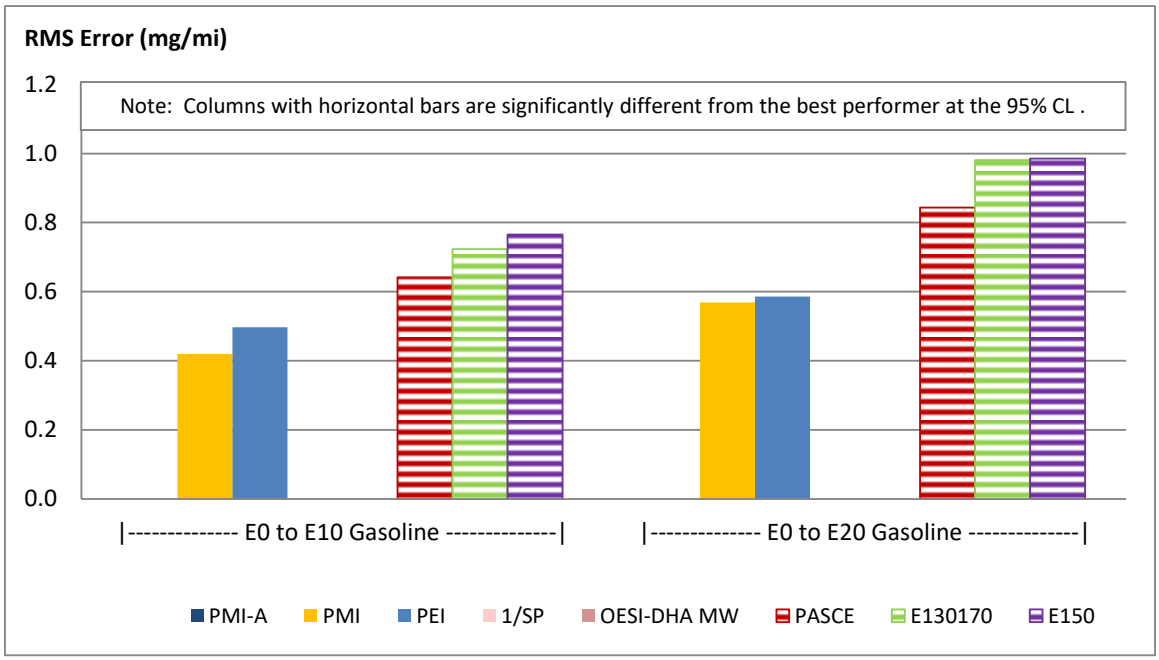
and that the effect is larger at higher PMI values than at lower, consistent with the expanding wedge that is observed.

7.1.2 RMS Error Performance Measure

Figure 7-3 shows the comparative performance of the PM indices for E0-E20 gasoline based on the RMS Error measure. Here, we see that the PMI is the best performer, although the PEI is also a high performer for both groups. These two speciation-based indices again outperform the correlation-based indices by clear-cut, if modest, margins. Given the relatively large sizes of the two EPAct groups (15 vehicles with 17 and 27 fuels), the statistical uncertainty in the performance measure is small. The 3 correlation-based indices are found to be statistically different (poorer) than the speciation-based indices. While the errors seem small in an absolute sense (<1 mg/mi), they are large in a relative sense for these PFI vehicles, given that their median Phase 1 PM emission levels are only 2.2 to 2.4 mg/mi for the two fuel groups, respectively.

Figure 7-3

Index Performance based on RMS Error: EPAct Fuels through E20



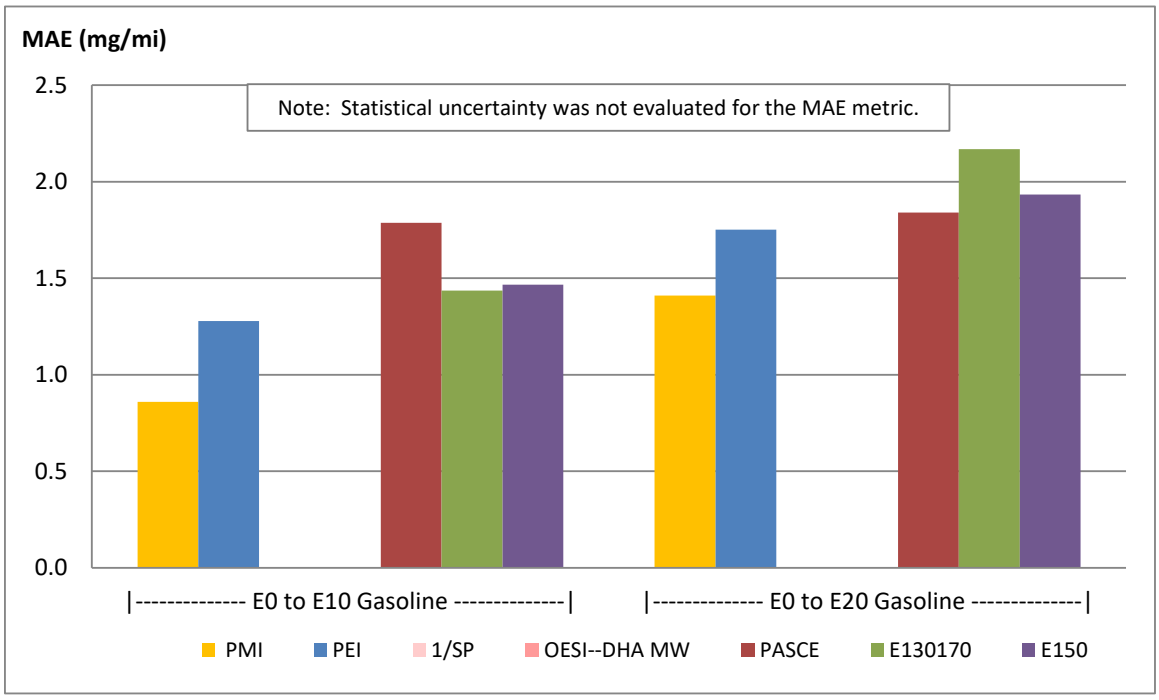
For all of the indices, the RMS Errors are somewhat larger for the group of E0-E20 gasolines; this result is consistent with the ethanol bias being larger than in the E0-E10 gasolines. As none of the indices account for the ethanol bias, it leads to larger errors in prediction in the E0-E20 group and to larger RMS Errors. In contrast, the Incremental R^2 measure was seen in the prior section to improve in the E0-E20 group. This measure, while related to the RMS Error, is one in which the errors in prediction are judged against the

overall variation of the data from its mean emission value. In many respects, the RMS Error is the clearest guide to the errors that will be encountered in using the indices.

7.1.3 MAE Performance Measure

Figure 7-4 shows the comparative performance of the PM indices based on the MAE measure. Here, the pattern of performance is nearly the same as for the RMS Error. The PMI is the best performer by a clear-cut margin over PEI, the other high performer, and by substantial margins over the correlation-based indices. The gap between the PMI and PEI is larger here than for the RMS Error. This is what one would expect if non-aromatic hydrocarbons play a larger role in the PM emissions for one fuel, as PEI accounts for only aromatics while the PMI (and PMI-A) account for all hydrocarbons in the speciation. The speciation-based indices would generally be preferred over the correlation-based indices whenever the worst-case error should be minimized.

Figure 7-4
Index Performance based on MAE: EPAct Fuels through E20



7.2 Summary of Ethanol Bias

This section summarizes the ethanol bias present in the five indices when applied to the EPAct fuels. Because of the limited ability to resolve the emission differences between E0 and E>0 gasolines at these low Phase 1 PM emission levels, only a brief discussion is presented.

Table 7-1 presents the characteristics of the ethanol bias for the EPAct fuels, including the size of the bias in the E0-E10 and E0-E20 fuel groups and the trend of the bias with index value. The average size of the bias (the average bias for E>0 gasoline minus the average bias for E0 gasoline) is larger for the E0-E20 fuels than for EPAct fuels up to E10. This is an indication that the effect of ethanol on PM emissions (at any given PM index value) increases in proportion to its usage in the fuel. Also, as often seen in the preceding figures, the ethanol effect is not observable at the very low PM emission levels of the low PMI fuels, indicated that the PM emissions effect may scale with the emissions level.

Table 7-1
Characteristics of the Ethanol Bias in PM Indices: EPAct Fuels

PM Index	Average Ethanol Bias (%)		Shape of Ethanol Bias
	E0-E10 Fuels	E0-E20 Fuels	
PMI	10%	22%	Expanding Wedge
PMI-A	(Not Evaluated)		
PEI	6%	17%	Expanding Wedge
1/SP	(Not Evaluated)		
OESI*-DHA MW	(Not Evaluated)		
PASCE	17%	25%	Expanding Wedge
E130-170	7%	12%	Narrowed Wedge
E150	3%	10%	Narrowed Wedge

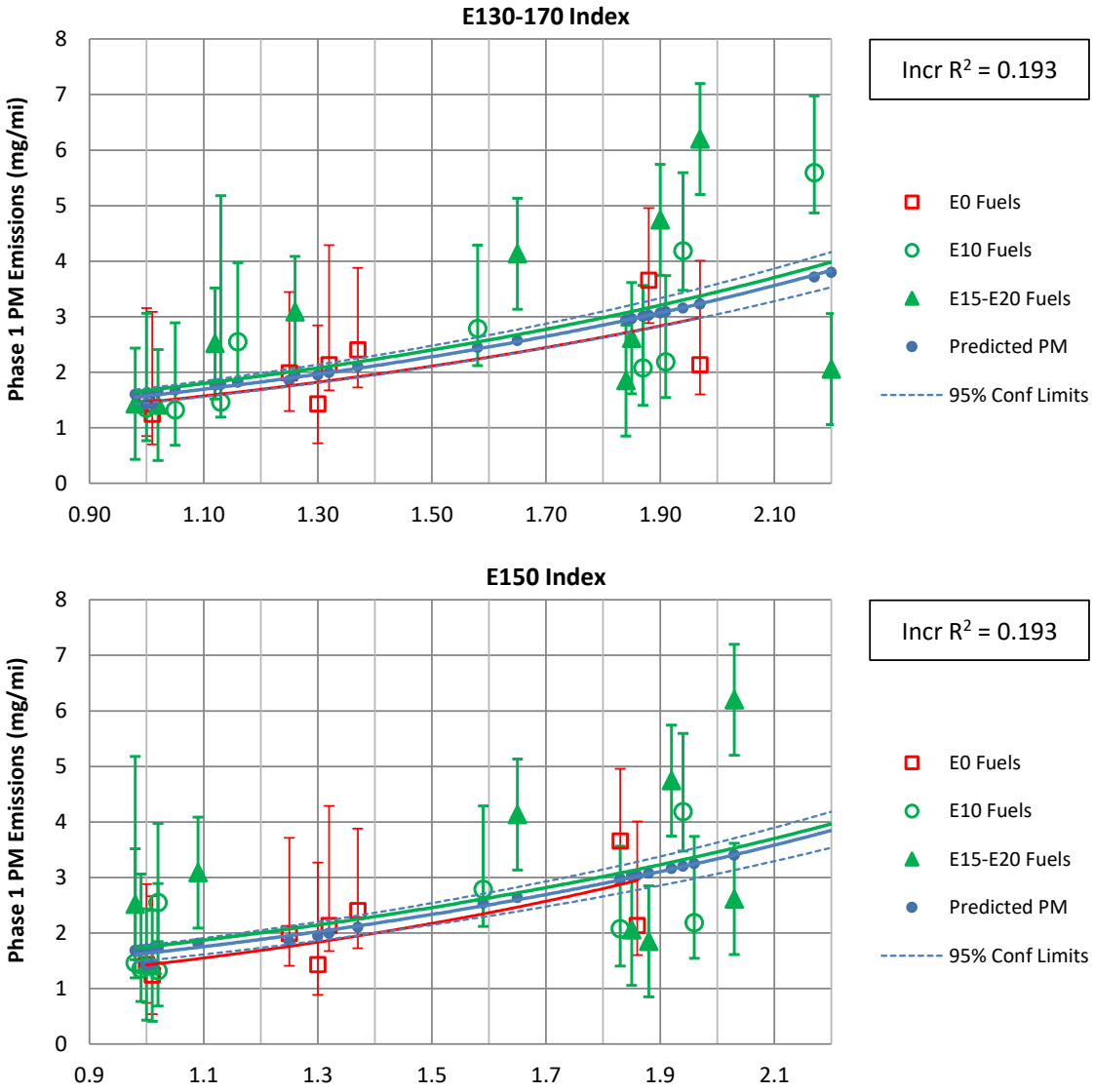
The E130-170 and E150 indices are the high performers in terms of ethanol bias, having the smallest biases by a large margin. PEI is next in line followed by the PMI and PASCE. The better performance of the two simplest indices on this basis is not without a cost, however, as it is comes with large RMS and MAE errors.

As the table also notes, the trend of the ethanol bias also differ from the other indices. The "expanding wedge" was seen for the PMI in the previous Figure 7-2. The performance graphs for the PEI and PASCE indices are quite similar. The vertical width of the "wedge" narrows considerably in the E130-170 index and essentially disappears for the E150 index.

Figure 7-5 shows the performance of the E130-170 and E150 indices, the two best performing indices with respect to the ethanol bias, when applied to the E0-E20 gasolines where the effect of ethanol on emissions should be largest. It is possible to resolve much narrowed wedges between the green trend line representing the E>0 gasolines and the red line for the E0 gasolines, but the scatter in the data is larger than the size of wedges. Given this, the two index methods can be described as nearly un-biased with respect to the effect of ethanol on PM emissions, but that comes at the expense of increased errors, particularly

for the fuels oxygenated with ethanol. For the latter, the range in worst-case errors for Phase 1 PM emissions –from the largest negative error (below the line) to the largest positive error (above the line) – is about 4 mg/mi or more than the 3 mg/mi range of the correlation line from the lowest to highest index values. When the indices are applied to a group of E>0 fuels, one would face committing an (unknown) error that could be as large as 2/3^{rds} the entire range from low to high PM emissions. Even if un-biased with respect to ethanol, one cannot count on good performance from the indices in their current forms.

Figure 7-5
Performance Graphs for the E130-170 and E150 Indices: E0-E20 Gasolines



7.3 Summary of Performance for E0-E10 and E0-E20 Gasoline

Table 7-2 compiles a summary of index performance for the EPAct study gasolines in an effort to find consensus. The high performing indices for each measure are indicated in the table by their rank.

PMI and PEI—the speciation-based indices—are the clear-cut and consistent best performers; one would expect PMI-A to join them if the index could be evaluated. These indices contain information on the hydrocarbon content of the fuels and carry information on the ethanol content through the reduction to index values when ethanol is present. For the EPAct fuels, where the design may have perturbed the correlations between composition and fuel properties, the correlation-based indices are disadvantaged in comparison.

Table 7-2
High Performing Indices for E0-E10 and E0-E20 Gasolines

		R^2	**RMS Error**	**MAE**
PMI-A	E0-E10	(Not Evaluated)		
	E0-E20			
PMI	E0-E10	#1	#1	#1
	E0-E20	#1	#2	#2
PEI	E0-E10	#1	#1	#1
	E0-E20	#1	#2	#2
1/SP	E0-E10	(Not Evaluated)		
	E0-E20			
OESI*—DHA MW	E0-E10	(Not Evaluated)		
	E0-E20			
PASCE	E0-E10			
	E0-E20			
E130-170	E0-E10			
	E0-E20			
E150	E0-E10			
	E0-E20			

8. SUMMARY AND CONCLUSIONS

This chapter presents a summary of findings and conclusions regarding the relative performance of the PM indices. Recommendations are given regarding the need for further development work on the indices and their potential applications in both research and operational settings.

8.1 Summary of Index Performance

8.1.1 E0 Gasolines

For the group of E0 gasolines tested in the three studies, all of the PM indices perform relatively well as predictors of emissions. The speciation-based indices (PMI, PMI-A, and PEI) are the consistently high performers across all three studies. For the two E-94 studies, they are joined by the PASCE and E130-170 indices as either high performers or close contenders. The E150 index and the two related indices that involve smoke point lag behind.

For the EPAct study, with its much lower Phase 1 PM emission levels of the PFI vehicles, only the PMI and PEI indices are high performers, but one would expect the PMI-A to join them if it could be evaluated for the EPAct fuels. These speciation-based indices have a performance advantage over the correlation-based indices throughout the comparisons. This is likely because the fuels embody distinctive and atypical relationships between hydrocarbon composition and bulk fuel properties due to the requirements of the EPAct experiment. The PMI and PEI indices have direct information on the hydrocarbon composition of the fuels that is not available to the correlation-based indices.

Correlation-based PM indices depend upon bulk fuel properties being reflections of and surrogates for the hydrocarbon composition of fuels. The indices are "tuned" during their development to the characteristics of the fuels on which they are based and, therefore, the correlations between hydrocarbon composition and the properties that are used as predictive variables. Correlation-based indices will, in general, not perform as well for new fuels when such fuels embody different relationships between composition and bulk properties than the fuels from which the indices were originally estimated.

8.1.2 E10 Gasolines

The group of E10 gasolines was not analyzed as a separate group. The best evidence for the performance of the PM indices for E10 fuels alone comes from the individual trend lines for E0 and E10 fuels in the analysis performed for the E0-E10 group. For the two E-94 studies, the individual E0 and E10 trend lines closely parallel each other and the overall

trend line for five high performing indices: PMI-A, PMI, PEI, PASCE and E130-170. One can infer from the parallelism that the five indices would perform as well for a group of E10 fuels as they do for the group of E0 fuels.

The E150 index trails in this comparison and the two related indices involving smoke point perform relatively poorly for ethanol-containing fuels. As was found in the group of E0 gasolines, only the three speciation-based indices (PMI, PMI-A and PEI) remain high performers for the EPAct fuels.

8.1.3 E0-E10 Gasolines

When the E0 and E10 gasolines are grouped, the three speciation-based indices prove again to be high performers and do so consistently across the three studies. For the two E-94 studies, the E130-170 and E150 indices join them as high performers or as close runner-ups, while the other correlation-based indices lag behind. For the EPAct study, only the speciation-based indices are high performers.

For a mixed group of fuels that vary in ethanol content, the three speciation-based indices carry information on the ethanol content of the fuels through the dilution of the gasoline hydrocarbons and the near-proportional reduction of the index values that results when ethanol is added. Among the correlation-based indices, the E130-170 index is unique in being based on distillation points that are diagnostic for both ethanol content (E130) and heavy hydrocarbon content (E170). It ranks among the high performers for the E0-E10 gasoline group for all three performance metrics along with the speciation-based indices. The other correlation-based indices respond to ethanol content in varying and less direct ways.

All of the indices exhibit a mathematical bias (a systematic error) with respect to the effect of ethanol on emissions, meaning that the E10 fuels lie consistently above the correlation line at higher emission levels, while the E0 fuels lie consistently below the line at lower emission levels. This is termed a "bias" because the indices generally under-predict the PM emissions of E10 fuels and over-predict the emissions of E0 fuels when the prediction model is based on data from both E0 and E10 fuels. The size of the vertical gap between the separate trend lines for E0 and E10 gasoline ranges from 10 to 37 percent of the median PM emissions level depending on the index. This finding is consistent with the two E-94 studies, which demonstrated that the presence of ethanol at the E10 level increases PM emissions above the emissions of E0 fuels at constant index values, and the EPAct study, which concluded that ethanol increased PM emissions for most, if not all, of the test vehicles for E10-E20 fuels.

For the E-94-2 study, the average emissions difference between E10 and E0 gasolines (holding the value of the index in question constant) is:

- Smallest for the E130-170 and E150 indices (10-11%)

- Of intermediate value for the three speciation-based indices and OESI*-DHA MW (18-20%)

- Largest for the PASCE and 1/SP indices (26% and 38%).

For the E-94-3 study, with its additional splash-blended E10 gasolines, the average emissions differences are larger for all indices, ranging from 23% to 37%.

For the EPAct study, the size of the ethanol bias is generally smaller in percentage terms and is difficult to observe given the much lower level of Phase 1 PM emissions of its PFI vehicles. Nevertheless, the fuels oxygenated with ethanol tend to have higher emissions than E0 fuels of comparable PM index values. The average difference is increased when E15-E20 fuels are included with E0-E10 fuels, which suggests that ethanol's impact on emissions depends on the ethanol.

All of the indices display an ethanol bias. Further development work is needed on the ethanol effect if the PM indices are to be used with mixed groups of fuels.

8.1.4 EPAct Gasolines Up to E20

The EPAct study tested gasolines ranging from E0 up to E20 in ethanol content. For both the E0-E10 and E0-E20 subgroups, the PMI and PEI indices prove to be best performers; one would expect the PMI-A to join them if it could be evaluated. Speciation-based indices have a clear advantage over the correlation-based indices for the EPAct fuels, for which the design and blending has likely caused their correlations between fuel composition and bulk properties to depart from those of the fuels from which the indices were developed. As noted above, the ethanol bias is larger in the group of E0-E20 fuels than in the E0-E10 group, suggesting that the effect of ethanol depends, at least approximately, on the ethanol content.

8.1.5 Enhanced versus PMI

The CRC AVFL-29 project developed a new DHA method that is able to speciate the heavy hydrocarbons in gasoline more completely. Its use with the Honda Equation to compute PMI-A substantially increases the PMI values of fuels that contain many heavy hydrocarbons. The AVFL-29 speciation clearly provides a more complete assessment of the individual hydrocarbon content (and presumably PM potential) for heavy gasolines. However, this advantage does not translate into improved predictive performance for PM emissions in this study. Throughout the comparisons, the PMI-A performs as well as the PMI in predicting the Phase 1 PM emissions of the SIDI vehicles in the two E-94 studies, but it does not demonstrate clearly better performance.

The analysis finds a very high level of statistical correlation between PMI and PMI-A values (see Figure 3-1). When using correlated variables, one variable (PMI-A) might be a more complete and reliable assessment of a characteristic, but it will *not* demonstrate notably better predictive performance than another variable (PMI) when the correlation

between the two is very high. This outcome is likely the result of how the E-94 fuels were blended to achieve High vs. Low PMI targets.

The test fuels in the E-94 studies were blended to target Low and High PMI levels, while remaining as representative of commercial gasoline as possible. But, as the experiment did not explore *different ways* in which high PMI values could be achieved, it is likely that only a very small number of heavy hydrocarbon streams were used (possibly only one or two). In this circumstance, the prevalence of the heavy hydrocarbons in a fuel (successfully identified in an ASTM DHA) should be strongly correlated to the prevalence of the heaviest hydrocarbons in the fuel (identified only in an AVFL-29 DHA). In this case, the PMI can predict PM emissions as well as the PMI-A, even if the PMI-A gives a more complete assessment of the hydrocarbon composition.

In other cases, such as fuels that are atypically high or low in the prevalence of the heaviest hydrocarbons, the PMI-A should be a better and more reliable predictor of emissions. This would be seen most easily in a dataset of fuels that had been purposely blended to explore the different ways in which heavy hydrocarbons occur in gasoline. Evidence of better performance by the PMI-A can be seen in some of the MAE performance comparisons for worst-case errors, but the apparent advantage is at most modest in size. The advantage should be larger for fuels that are atypical in the prevalence of the heaviest hydrocarbons.

8.2 Improvements to the PM Indices

An effort was made in the analysis to identify fuel variables whose inclusion in the index might improve the predictive performance of the PM indices for emissions. The outcome of this work does not, and cannot, provide definitive answers to the question, but it should provide some guidance for future development work.

The analysis indicates that PMI, PMI-A and PEI would not benefit from the inclusion of additional variables, except with regard to the emissions effect of ethanol. These indices are based on speciation of the hydrocarbon content of fuels—involving all hydrocarbon compounds in the PMIs, but only the aromatic hydrocarbons in PEI—and apparently gain little from the addition of the bulk fuel properties considered.

For the E0 gasolines, fuel density is the main variable identified as having the potential to improve the performance of the correlation-based indices. For the E0-E10 gasolines, after ethanol was included as a predictive variable, a number of additional fuel properties were identified as potentially of benefit, including fuel density, aromatics and olefins content, and subsets of light and heavy aromatics content. The identified variables are logical choices because their inclusion could help differentiate between light and heavy gasolines that may otherwise be similar in the bulk fuel properties used as predictors.

8.3 Representation of the Ethanol Effect on Emissions

All of the indices are affected by a bias in estimating the Phase 1 PM emissions of fuels oxygenated with ethanol. The indices under-estimate the emissions of E10 fuels and over-

estimate the emissions of E0 fuels by amounts that range from 10 to more than 30 percent depending on the index. If applied to a mixed group of E0-E10 fuels, the indices may mis-rank the fuels for emissions, tending to favor E10 over E0 fuels. Work is needed to correct this bias if the indices are to be used in daily operations with mixed groups of fuels.

Why ethanol has an adverse effect on PM emissions for E0-E20 fuels is not fully understood. The combustion of ethanol should not contribute substantially to particle formation, as ethanol is a saturated hydrocarbon compound with a high vapor pressure and a correspondingly small "i-term" in the Honda Equation. Its presence dilutes the gasoline hydrocarbons and the PMI value of a fuel, but its presence also leads to higher PM emissions than would be expected from the PMI of the oxygenated fuel.

One hypothesis is that ethanol's high heat of vaporization is responsible for its impact on PM emissions. For example, Butler 2015 found that there was a reinforcing interaction of ethanol on PM emissions in 10 of the 15 test vehicles from the EPAct study and went on to say that their results were consistent with statements in previous studies "associating ethanol's higher heat of vaporization with a cooling effect that has the potential to hinder fuel evaporation and lead to increased PM emissions." Barrientos 2016 explored heat of vaporization and net heating value as variables that might improve the predictive performance of the OESI index for emissions. Chen 2018 and Burke 2018 give more recent examinations of the heat of vaporization effect of ethanol. In addition to past work, the CRC E-129 project, which is testing three different oxygenates at two different oxygen levels, may shed light on the ethanol effect and whether it extends to other oxygenated compounds.

The predictive performance of all indices can be improved once the mechanism of the ethanol effect is better understood. Until such time, the indices could be improved by work to incorporate empirical adjustments for ethanol content, recognizing possible effects due to engine and operating condition differences.

8.4 Use of the Indices in Research and Operations

The three speciation-based indices—PMI, PMI-A and PEI—are the consistent high performers across the three studies and clearly outperform the correlation-based indices for the EPAct fuels. The speciation-based indices have detailed information regarding the hydrocarbon composition of fuels. For this reason, they may be able to perform well across a wide range of gasoline fuels. In purely research settings, PMI and PMI-A may be preferred because they consider the contributions from all hydrocarbon compounds, not just the aromatic hydrocarbons considered in PEI. PMI-A may be better predictor of emissions in studies of fuels containing the heaviest hydrocarbons and whenever atypical fuels are involved.

These three indices could be used in an operational setting if the needed hydrocarbon speciation is available. As these indices demonstrate comparable levels of performance for fuels of the same oxygenation level (i.e., E0 or E10), the preference for one index over

another is more likely to be determined by one's capability and preference for hydrocarbon speciation, rather than the performance of the index itself.

In narrow circumstances, specifically where the fuels are generally similar to the E-94 fuels and only a single level of ethanol is involved (E0 or E10), the PASCE and E130-170 indices might achieve similar levels of performance to the three speciation-based indices. E130-170 could also be used if only the distillation curve were known and PASCE could be used when the curve and the C/H ratio is known. The related 1/SP and OESI*-DHA MW indices do not achieve the same level of performance as the other indices and prove to have difficulty representing PM emissions for mixed groups of E0-E10 fuels.

All of the indices require further development to account in an unbiased manner for the PM emissions potential of both E0 and E10-E20 gasoline. The correlation-based indices may also benefit from inclusion of additional variables such as fuel density. Until this additional development work has been done, the best performing correlation-based indices could be used in daily operations only where attention is given to "tuning" them to the specific group of fuels that will be encountered. In essence, the result of such "tuning" would be to create a tailored index for an individual facility or operation.

9. REFERENCES

Aikawa, K., Sakurai, T., Jetter, J. (2010). *Development of a Predictive Model for Gasoline Vehicle Particulate Matter Emissions*. SAE Technical Paper 2010-01-2115.

Barrientos, E., M.A. Lapuerta, A.L. Boehman (2013). *Group additivity in soot formation for the example of C-5 oxygenated hydrocarbon fuels*. Combustion and Flame, Volume 160, Issue 8, August 2013, pages 1484-1498.

Barrientos, E., Anderson, J., Mariq, M., Boehman, A. (2016). *Particulate Matter Indices using Fuel Smoke Point for Vehicle Emissions with Gasoline, Ethanol Blends, and Butanol Blends*. Volume 167, Combustion and Flame, May 2016, pages 308-319.

Burke, S., Rhoads, R., Ratcliff, M., McCormick, R., Windom, B. (2018). *Measured and Predicted Vapor Liquid Equilibrium of Ethanol-Gasoline Fuels with Insight on the Influence of Azeotrope Interactions on Aromatics Species Enrichment and Particulate Matter Formation in Spark Ignition Engines*. April 2018. NREL/CP-5400-71336.

Butler, A., Sobotowski, R., Hoffman, G., and Machiele, P. (2015). *Influence of Fuel PM Index and Ethanol Content on Particulate Emissions from Light-Duty Gasoline Vehicles*. SAE Technical Paper 2015-01-1072, 2015.

Calcote, H.F., D.M. Manos (1983). *Effect of Molecular Structure on Incipient Soot Formation*. Combustion and Flame, Volume 49, Issues 1–3, January 1983, pages 289-304.

Chapman, E., Winston-Galant, M., Geng, P., Latigo, R., Boehman, A. (2016). *Alternative Fuel Property Correlations to the Honda Particulate Matter Index (PMI)*. SAE Technical Paper 2016-01-2250.

Chapman, E., Geng, P., Winston-Galant, M. (2018). *Release of the GM Particulate Emissions Index (PEI) Equation Externally*. Communication with CRC, June 19, 2018.

Chen, Yu. (2018). *Effects of Ethanol Evaporative Cooling on Particulate Number Emissions in GDI Engines*. SAE Technical Paper 2018-01-0360.

Cho, J., Si, W., Jang, W., Jin, D., Myung, C., Park, S. (2015). *Impact of intermediate ethanol blends on particular matter emission from a spark ignition direct injection (SIDI) engine*. Applied Energy, 10.1016/j.apenergy.2015.08.010.

CRC AVFL-29 (2018), *Enhanced Speciation of Gasoline*. CRC Project AVFL-29. June 2018.

CRC E-94-2 (2017), *Evaluation and Investigation of Fuel Effects on Gaseous and Particulate Emissions on SIDI In-Use Vehicles.* CRC Project E-94-2. March 2017.

CRC E-94-3 (2018), *Impacts of Splash-Blending on Particulate Emissions for SIDI Engines*. CRC Project E-94-3. June 2018.

EPAct (2010), *EPAct/V2/E-89 Fuel Blending, Final Report.* Whitney, K., Southwest Research Institute, September 2010.

EPAct (2013a), *EPAct/V2/E-89: Assessing the Effect of Five Gasoline Properties on Exhaust Emissions from Light-Duty Vehicles Certified to Tier 2 Standards, Analysis of Data from EPAct Phase 3*. EPA-420-R-13-002. April 2013.

EPAct (2013b), *EPAct/V2/E-89: Assessing the Effect of Five Gasoline Properties on Exhaust Emissions from Light-Duty Vehicles Certified to Tier 2 Standards, Final Report on Program Design and Data Collection*. EPA-420-R-13-004. April 2013.

Moriya, H. (2016). *Fuel Property Influence on Exhaust Emissions: Simplified PM Index*. SAE-China Congress & Exhibition, Shanghai, October 26, 2016.

Sakai, S. and D. Rothamer (2019). *Impact of ethanol blending on particulate emissions from a spark-ignition direct-injection engine*. Fuel, Volume 236, 15 January 2019, pages 1548-1558.

APPENDIX A

PERFORMANCE RESULTS FOR E0 GASOLINE

Table A-1
Summary of PMI Performance for E0 Gasoline

	E-94-2	E-94-3	EPAct
Median Phase 1 PM Emissions (mg/mi)	33.0	29.3	2.8
Incremental R^2	0.850	0.785	0.297
RMS Error	2%	2%	16%
mg/mi	0.8	0.7	0.5
MAE Error	3%	3%	31%
mg/mi	1.1	0.8	0.8
PM index of the fuel	2.49	1.40	1.30
Fuel variables adding explanatory power ($p \leq 0.01$)	None	None	AKI

Table A-2
Summary of PMI-A Performance for E0 Gasoline

	E-94-2	E-94-3	EPAct
Median Phase 1 PM Emissions (mg/mi)	33.0	29.3	2.8
Incremental R^2	0.849	0.788	(Not Evaluated)
RMS Error	3%	3%	
mg/mi	0.9	0.8	
MAE Error	3%	3%	
mg/mi	1.0	1.0	
PM index of the fuel	1.32	2.31	
Fuel variables adding explanatory power ($p \leq 0.01$)	None	None	

Table A-3
Summary of PEI Performance for E0 Gasoline

	E-94-2	E-94-3	EPAct
Median Phase 1 PM Emissions (mg/mi)	33.0	29.3	2.8
Incremental R^2	0.838	0.787	0.288
RMS Error	5%	2%	17%
mg/mi	1.6	0.5	0.5
MAE Error	6%	2%	36%
mg/mi	2.1	0.6	1.0
PM index of the fuel	0.93	0.93	0.75
Fuel variables adding explanatory power ($p \leq 0.01$)	None	None	T60

Table A-4
Summary of 1/SP Performance for E0 Gasoline

	E-94-2	E-94-3	EPAct
Median Phase 1 PM Emissions (mg/mi)	33.0	29.3	2.8
Incremental R^2	0.634	0.633	(Not Evaluated)
RMS Error	18%	14%	
mg/mi	6.0	4.0	
MAE Error	29%	20%	
mg/mi	9.6	5.8	
PM index of the fuel	0.053	0.053	
Fuel variables adding explanatory power ($p \leq 0.01$)	Density	Density	

Table A-5
Summary of OESI*-DHA MW Performance for E0 Gasoline

	E-94-2	E-94-3	EPAct
Median Phase 1 PM Emissions (mg/mi)	33.0	29.3	2.8
Incremental R^2	0.716	0.685	(Not Evaluated)
RMS Error	15%	11%	
mg/mi	4.9	3.3	
MAE Error	22%	16%	
mg/mi	7.4	4.6	
PM index of the fuel	0.555	0.613	
Fuel variables adding explanatory power ($p \leq 0.01$)	Density	None	

Table A-6
Summary of PASCE Performance for E0 Gasoline

	E-94-2	E-94-3	EPAct
Median Phase 1 PM Emissions (mg/mi)	33.0	29.3	2.8
Incremental R^2	0.838	0.786	0.197
RMS Error	5%	2%	27%
mg/mi	1.6	0.6	0.8
MAE Error	6%	2%	60%
mg/mi	1.9	0.7	1.7
PM index of the fuel	1.09	1.87	1.68
Fuel variables adding explanatory power ($p \leq 0.01$)	None	None	Density

Table A-7
Summary of E130-170 Performance for E0 Gasoline

	E-94-2	E-94-3	EPAct
Median Phase 1 PM Emissions (mg/mi)	33.0	29.3	2.8
Incremental R^2	0.846	0.785	0.148
RMS Error	3%	2%	32%
mg/mi	1.1	0.7	0.9
MAE Error	4%	4%	56%
mg/mi	1.3	1.0	1.5
PM index of the fuel	1.16	1.71	1.60
Fuel variables adding explanatory power ($p \leq 0.01$)	None	None	Density

Table A-8
Summary of PASCE Performance for E0 Gasoline

	E-94-2	E-94-3	EPAct
Median Phase 1 PM Emissions (mg/mi)	33.0	29.3	2.8
Incremental R^2	0.795	0.749	0.104
RMS Error	10%	7%	35%
mg/mi	3.1	2.0	1.0
MAE Error	13%	10%	59%
mg/mi	4.4	2.8	1.6
PM index of the fuel	1.14	1.52	1.57
Fuel variables adding explanatory power ($p \leq 0.01$)	Density	None	Density

Figure A-1
Predictive Performance of PMI for E0 Gasoline

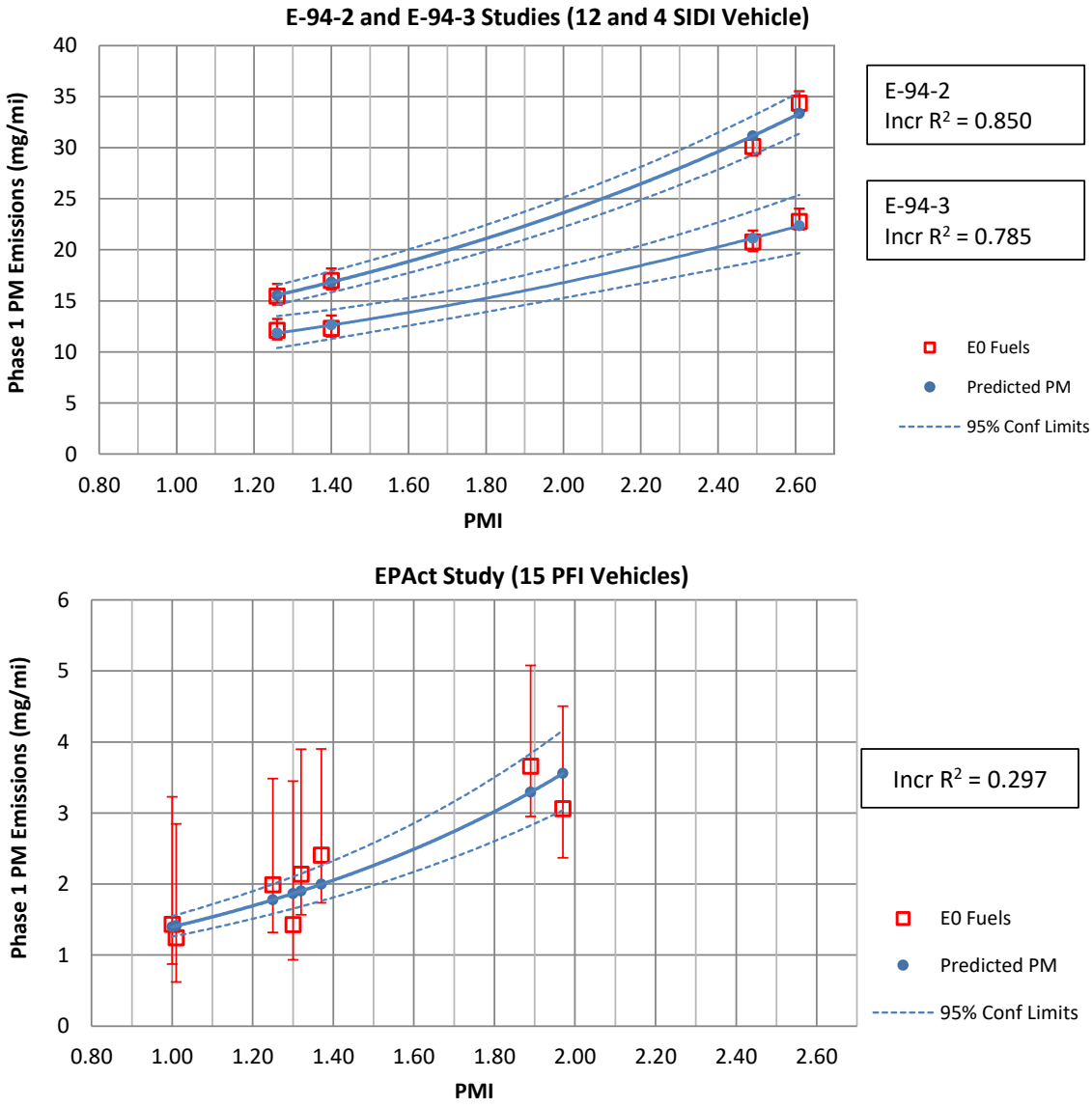


Figure A-2
Predictive Performance of PMI-A for E0 Gasoline

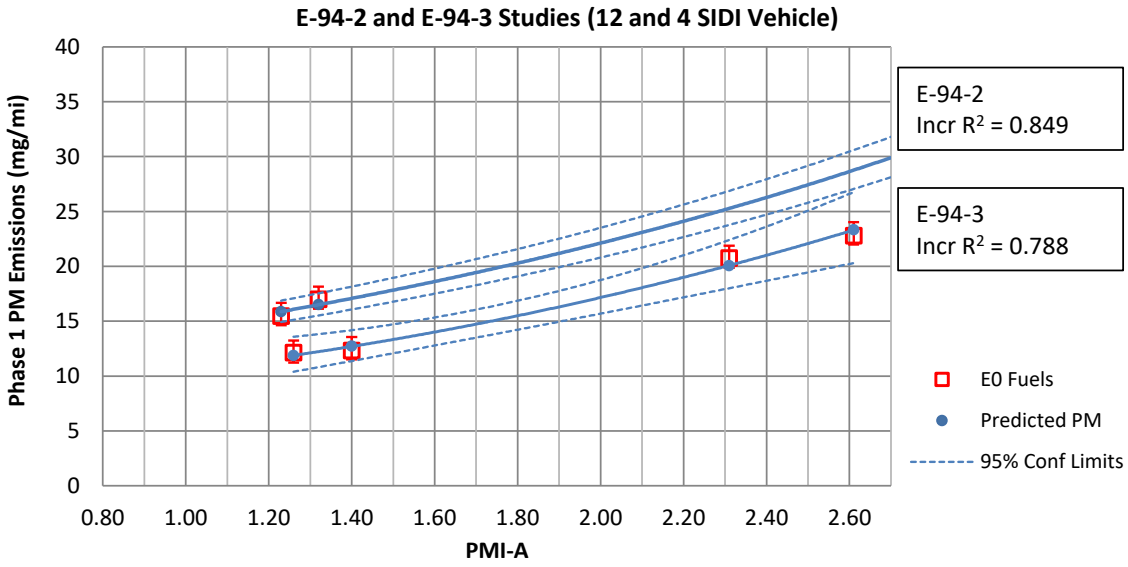


Figure A-3
Predictive Performance of PEI for E0 Gasoline

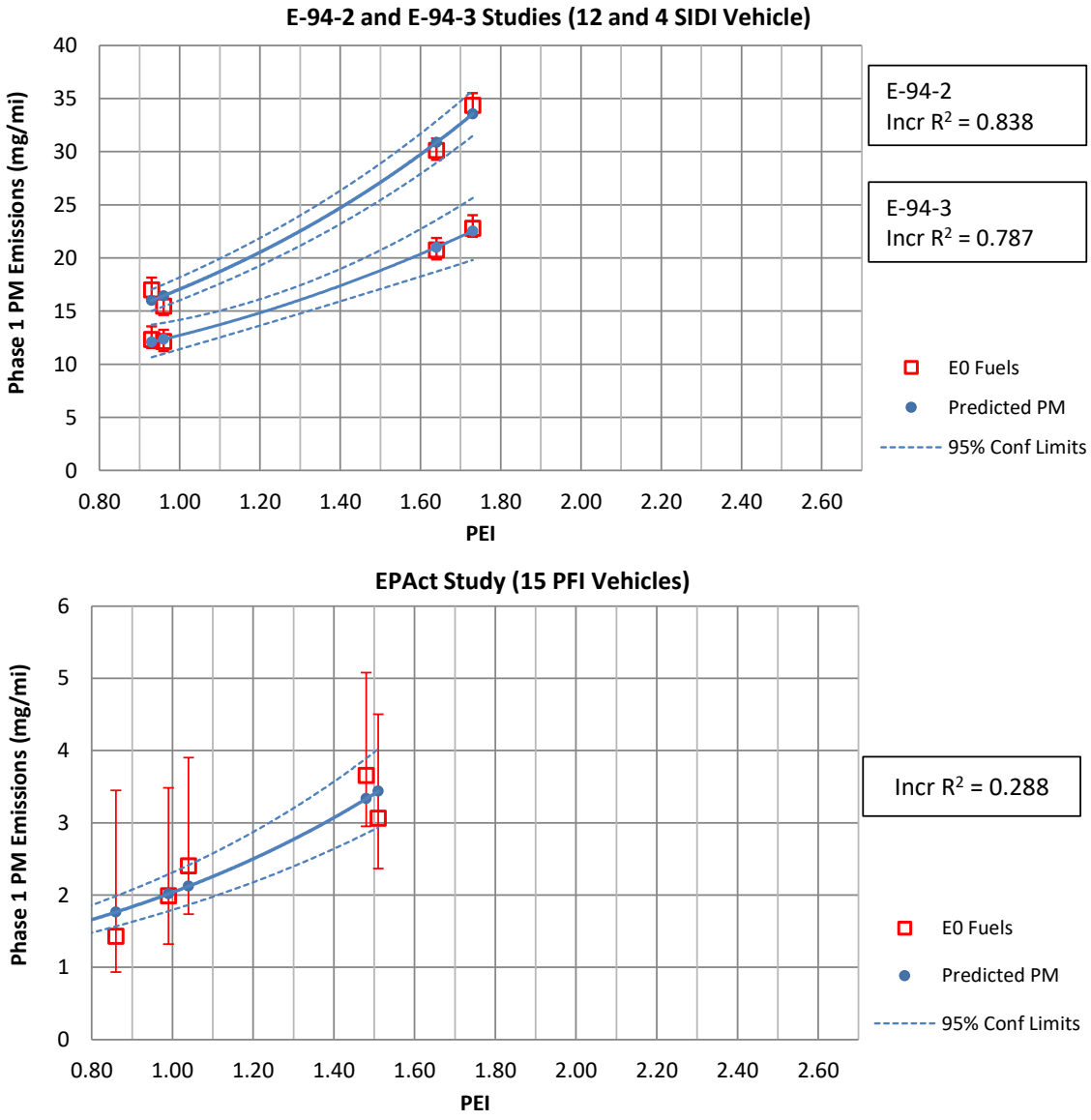


Figure A-4
Predictive Performance of 1/SP for E0 Gasoline

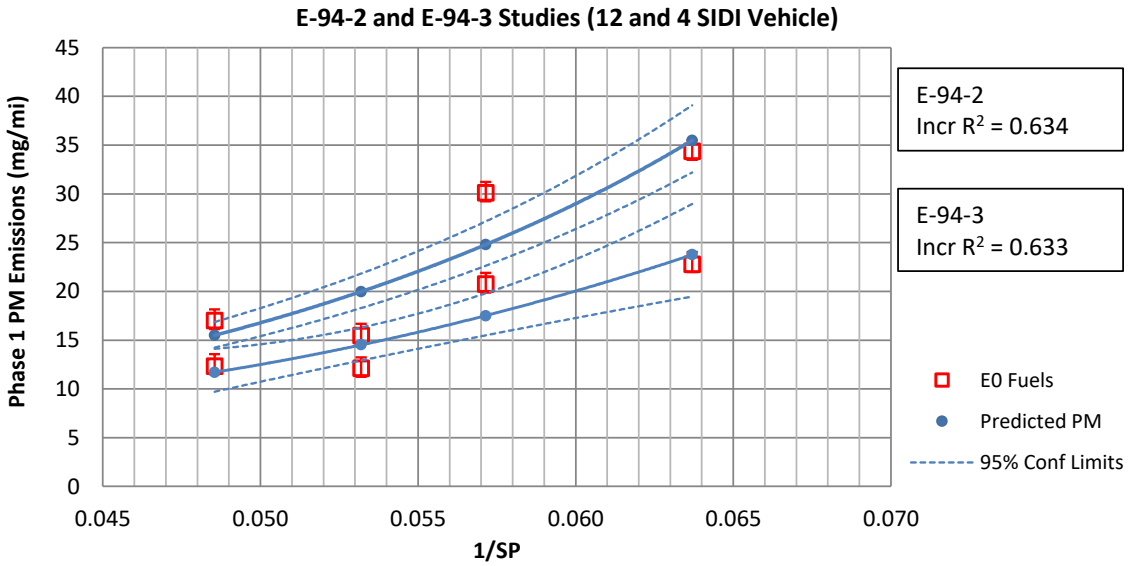


Figure A-5
Predictive Performance of OESI*-DHA MW for E0 Gasoline

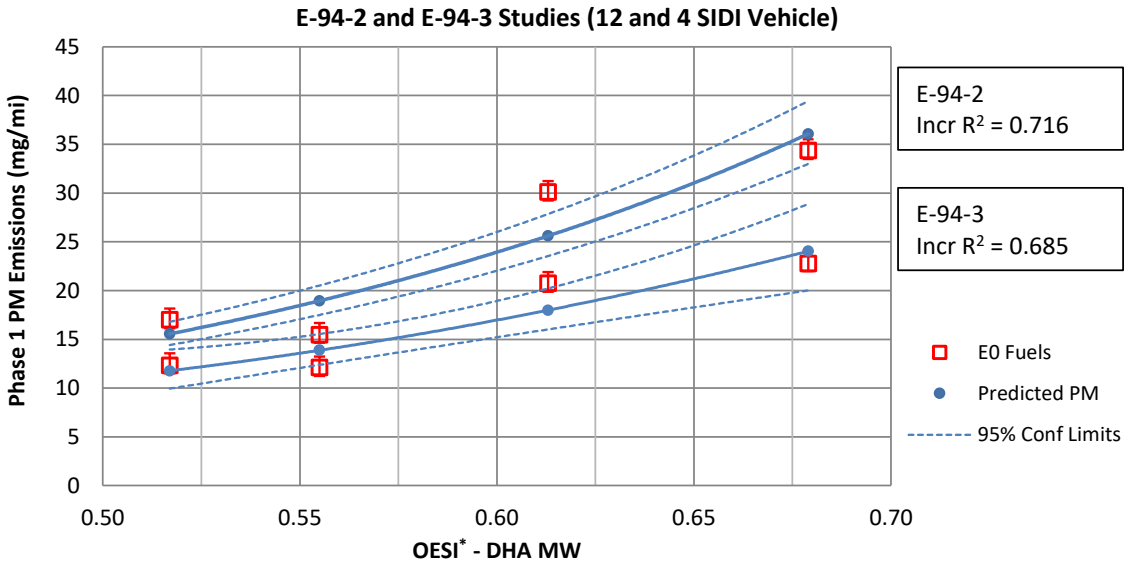
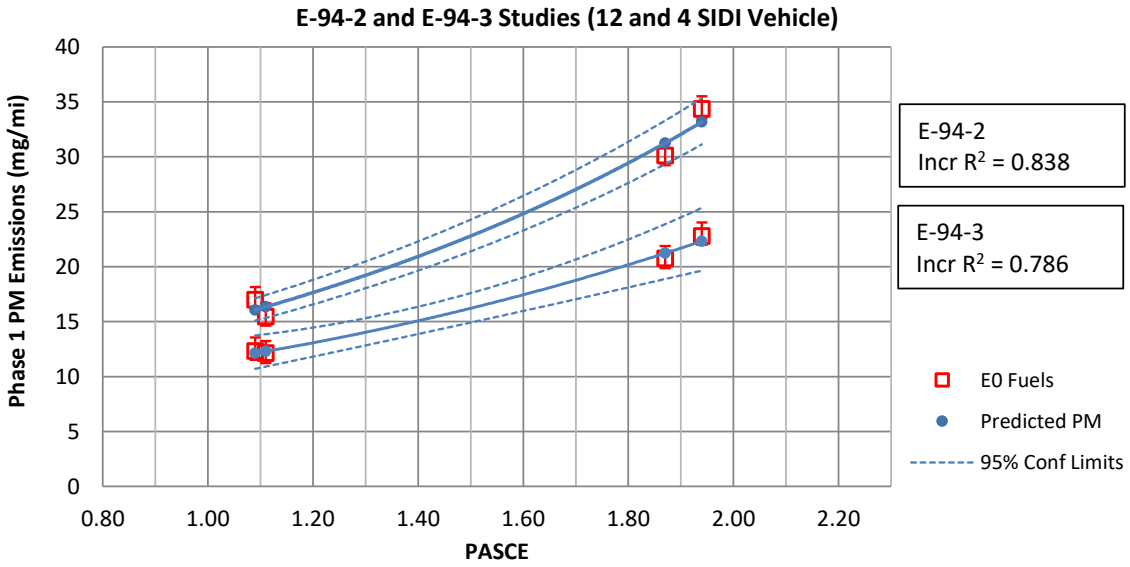


Figure A-6
Predictive Performance of PASCE for E0 Gasoline



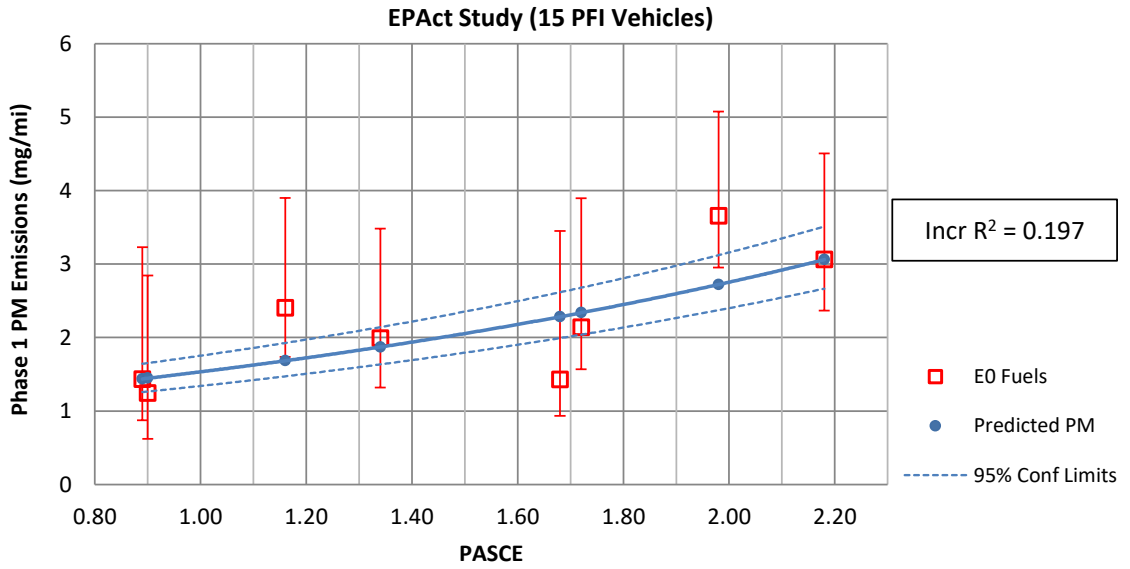
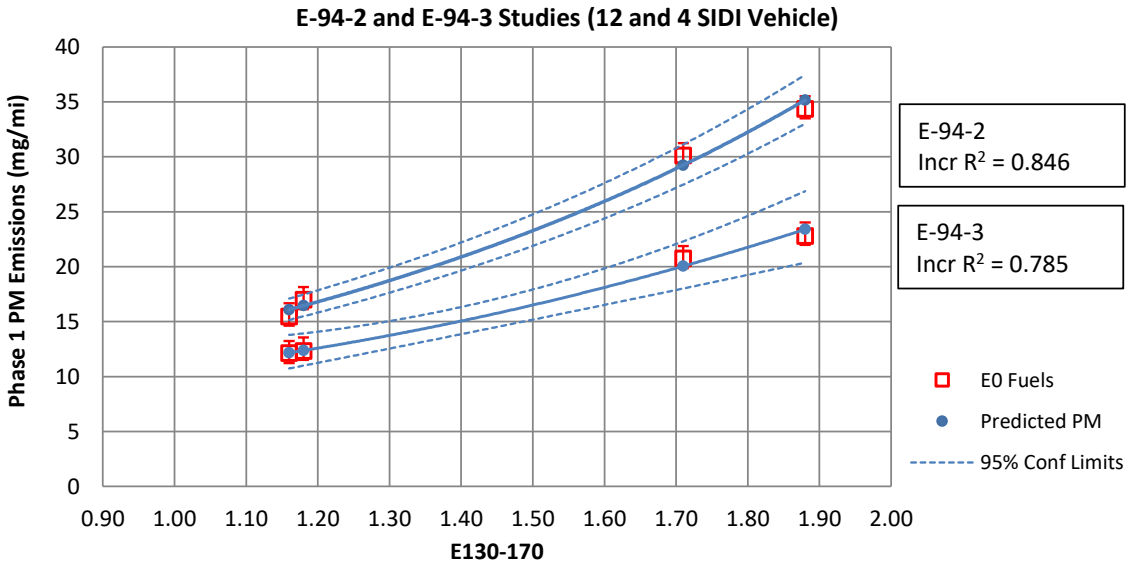


Figure A-7
Predictive Performance of E130-170 for E0 Gasoline



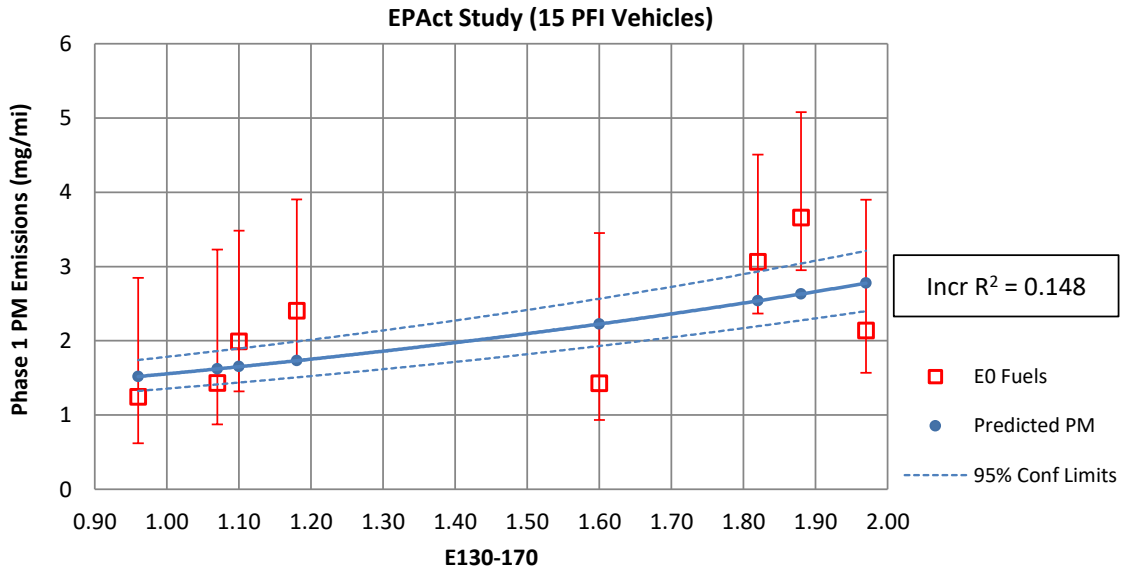
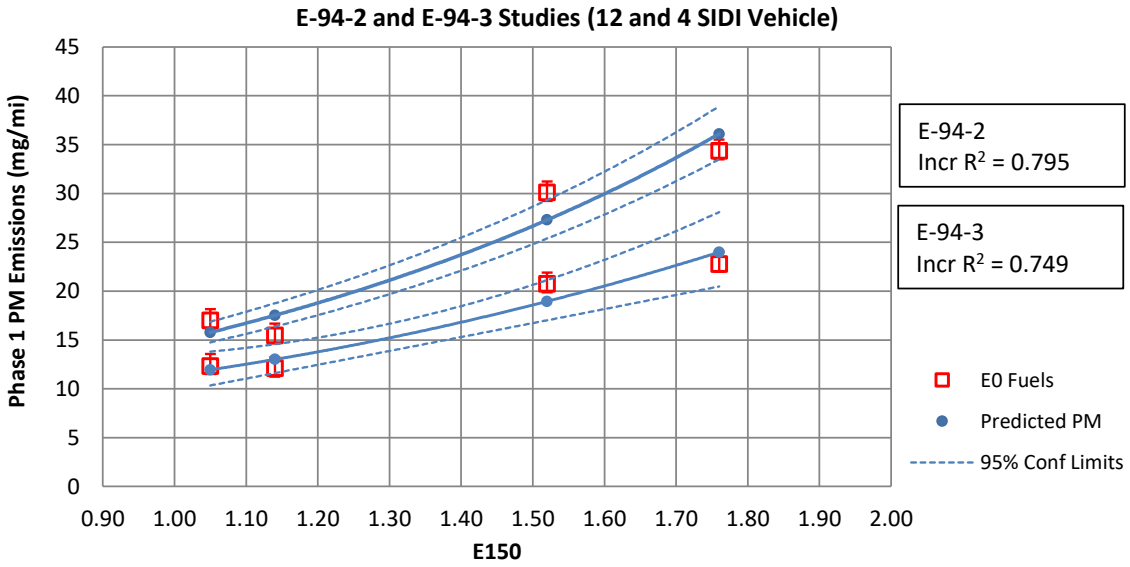
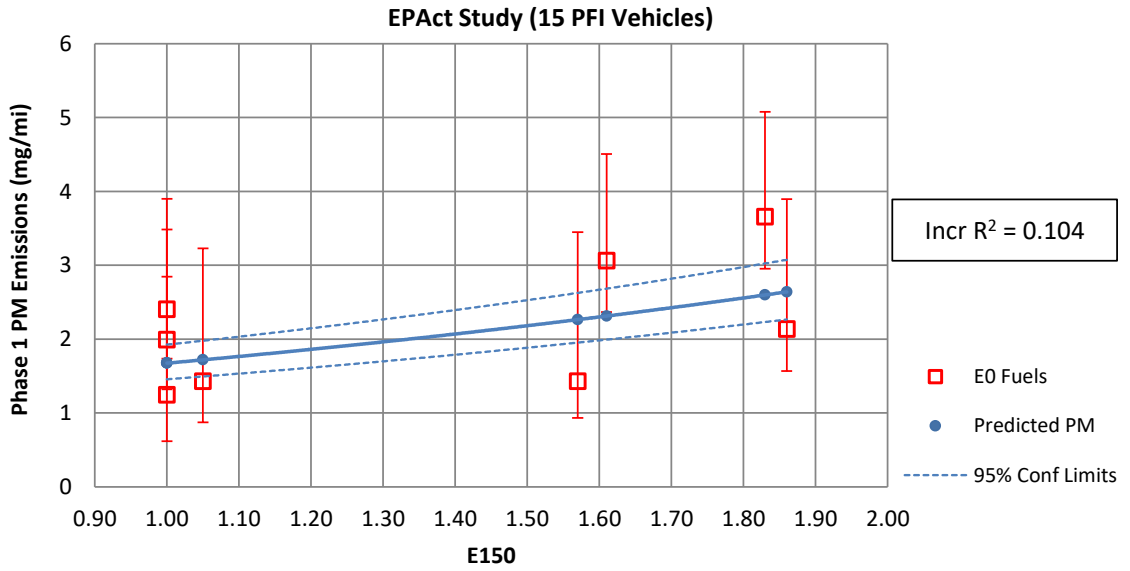


Figure A-8
Predictive Performance of E150 for E0 Gasoline





APPENDIX B

PERFORMANCE RESULTS FOR E0-E10 GASOLINE

Table B-1
Summary of PMI Performance for E0-E10 Gasolines

	E-94-2	E-94-3	EPAct
Median Phase 1 PM Emissions (mg/mi)	25.3	19.1	2.4
Incremental R^2	0.771	0.614	0.324
RMS Error	11%	14%	19%
mg/mi	2.9	2.7	0.4
MAE Error	23%	26%	39%
mg/mi	5.7	5.0	0.9
PM index of the fuel	1.28	1.40	1.34
EtOH Bias: Avg. Predictive Error (E>0 minus E0 fuels)	19%	25%	10%
For Low PMI Fuels	24%	29%	6%
For High PMI Fuels	11%	21%	24%
Fuel variables adding explanatory power ($p \leq 0.01$)			
When EtOH term is required	Benzene	None	None

Table B-2
Summary of PMI-A Performance for E0-E10 Gasolines

	E-94-2	E-94-3	EPAct
Median Phase 1 PM Emissions (mg/mi)	25.3	19.1	2.4
Incremental R^2	0.777	0.618	(Not Evaluated)
RMS Error	11%	14%	
mg/mi	2.8	2.6	
MAE Error	22%	23%	
mg/mi	5.5	4.3	
PM index of the fuel	1.25	1.32	
EtOH Bias: Avg. Predictive Error (E>0 minus E0 fuels)	18%	24%	
For Low PMI Fuels	22%	27%	
For High PMI Fuels	11%	22%	
Fuel variables adding explanatory power ($p \leq 0.01$)			
When EtOH term is required	Benzene	None	

Table B-3
Summary of PEI Performance for E0-E10 Gasolines

	E-94-2	E-94-3	EPAct
Median Phase 1 PM Emissions (mg/mi)	25.3	19.1	2.4
Incremental R^2	0.787	0.610	0.300
RMS Error	10%	14%	23%
mg/mi	2.6	2.7	0.5
MAE Error	18%	26%	59%
mg/mi	4.6	4.9	1.3
PM index of the fuel	0.96	0.96	1.09
EtOH Bias: Avg. Predictive Error (E>0 minus E0 fuels)	18%	26%	6%
For Low PMI Fuels	22%	28%	2%
For High PMI Fuels	11%	23%	21%
Fuel variables adding explanatory power ($p \leq 0.01$) When EtOH term is required	None	None	T60

Table B-4
Summary of 1/SP Performance for E0-E10 Gasolines

	E-94-2	E-94-3	EPAct
Median Phase 1 PM Emissions (mg/mi)	25.3	19.1	2.4
Incremental R^2	25.3	19.1	(Not Evaluated)
RMS Error			
mg/mi	0.554	0.249	
MAE Error	22%	26%	
mg/mi	5.7	5.0	
PM index of the fuel	44%	56%	
EtOH Bias: Avg. Predictive Error (E>0 minus E0 fuels)	11.2	10.7	
For Low PMI Fuels	0.06	0.05	
For High PMI Fuels	20%	32%	
Fuel variables adding explanatory power ($p \leq 0.01$) When EtOH term is required	AKI, RVP, Aromatics, Density	< C10 Aromatics	

Table B-5
Summary of OESI*-DHA MW Performance for E0-E10 Gasolines

	E-94-2	E-94-3	EPAct
Median Phase 1 PM Emissions (mg/mi)	25.3	19.1	2.4
Incremental R^2	0.116	0.024	(Not Evaluated)
RMS Error	37%	32%	
mg/mi	9.4	6.1	
MAE Error	69%	63%	
mg/mi	17.4	12.1	
PM index of the fuel	0.57	0.57	
EtOH Bias: Avg. Predictive Error (E>0 minus E0 fuels)	38%	30%	
For Low PMI Fuels	28%	26%	
For High PMI Fuels	52%	35%	
Fuel variables adding explanatory power ($p \leq 0.01$) When EtOH term is required	C10 Aromatics	C10 Aromatics	

Table B-6
Summary of PASCE Performance for E0-E10 Gasolines

	E-94-2	E-94-3	EPAct
Median Phase 1 PM Emissions (mg/mi)	25.3	19.1	2.4
Incremental R^2	0.712	0.535	0.249
RMS Error	15%	17%	29%
mg/mi	3.8	3.3	0.6
MAE Error	24%	32%	82%
mg/mi	6.0	6.1	1.8
PM index of the fuel	1.11	1.11	1.68
EtOH Bias: Avg. Predictive Error (E>0 minus E0 fuels)	26%	31%	17%
For Low PMI Fuels	29%	33%	10%
For High PMI Fuels	20%	28%	46%
Fuel variables adding explanatory power ($p \leq 0.01$) When EtOH term is required	None	None	DVPE, Olefins

Table B-7
Summary of E13-170 Performance for E0-E10 Gasolines

	E-94-2	E-94-3	EPAct
Median Phase 1 PM Emissions (mg/mi)	25.3	19.1	2.4
Incremental R^2	0.805	0.581	0.215
RMS Error	9%	15%	33%
mg/mi	2.2	2.9	0.7
MAE Error	17%	24%	66%
mg/mi	4.2	4.6	1.4
PM index of the fuel	1.70	1.53	1.60
EtOH Bias: Avg. Predictive Error (E>0 minus E0 fuels)	11%	24%	7%
For Low PMI Fuels	14%	25%	-4%
For High PMI Fuels	8%	23%	52%
Fuel variables adding explanatory power ($p \leq 0.01$)			
When EtOH term is required	Olefins	None	Aromatics

Table B-8
Summary of E150 Performance for E0-E10 Gasolines

	E-94-2	E-94-3	EPAct
Median Phase 1 PM Emissions (mg/mi)	25.3	19.1	2.4
Incremental R^2	0.766	0.600	0.197
RMS Error	12%	15%	35%
mg/mi	2.9	2.8	0.8
MAE Error	23%	30%	67%
mg/mi	5.9	5.8	1.5
PM index of the fuel	1.14	1.14	1.57
EtOH Bias: Avg. Predictive Error (E>0 minus E0 fuels)	10%	22%	3%
For Low PMI Fuels	18%	27%	3%
For High PMI Fuels	3%	18%	5%
Fuel variables adding explanatory power ($p \leq 0.01$)			
When EtOH term is required	T60, Density	None	Aromatics

Figure B-1
Predictive Performance of PMI for E0-E10 Gasolines

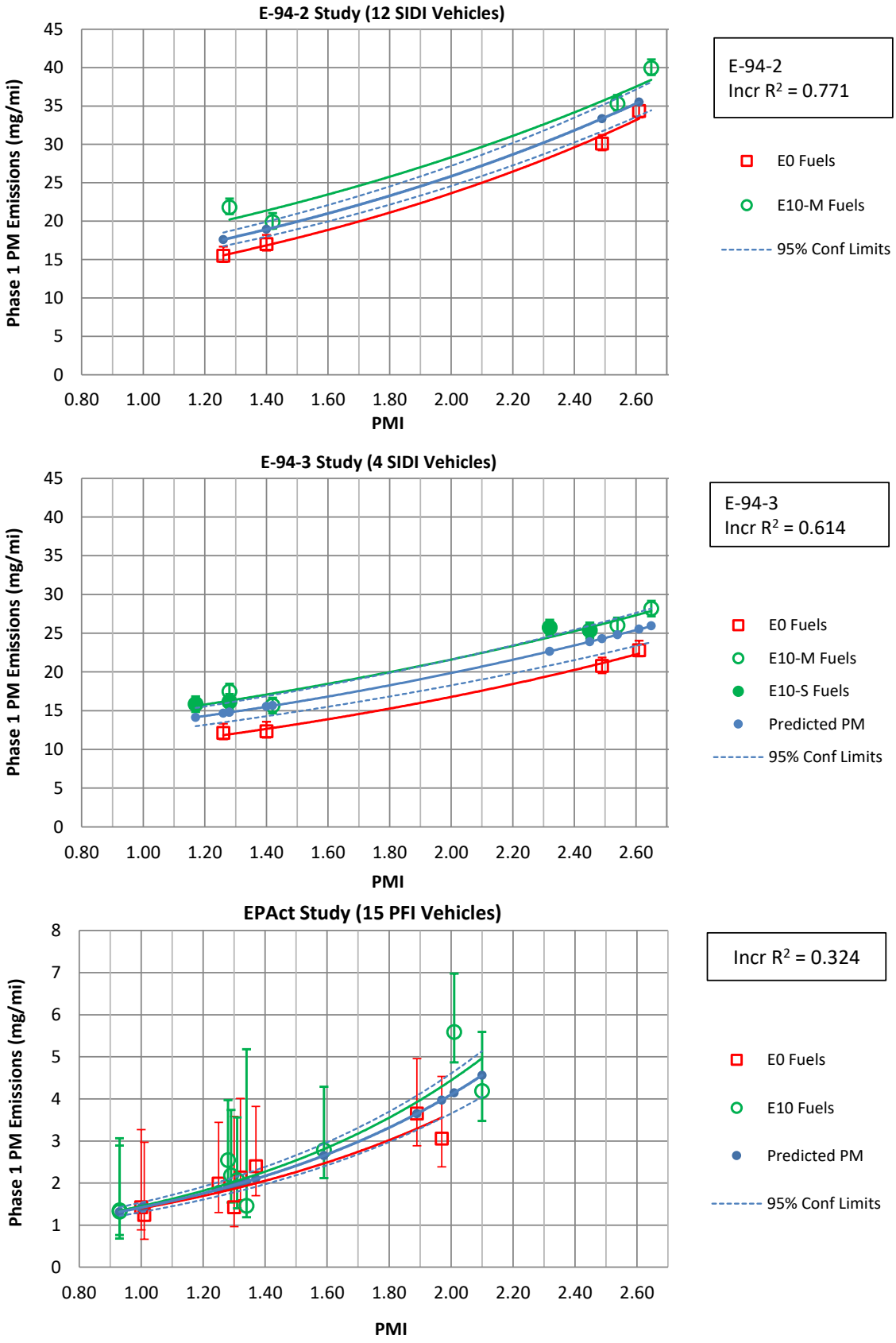


Figure B-2
Predictive Performance of PMI-A for E0-E10 Gasoline

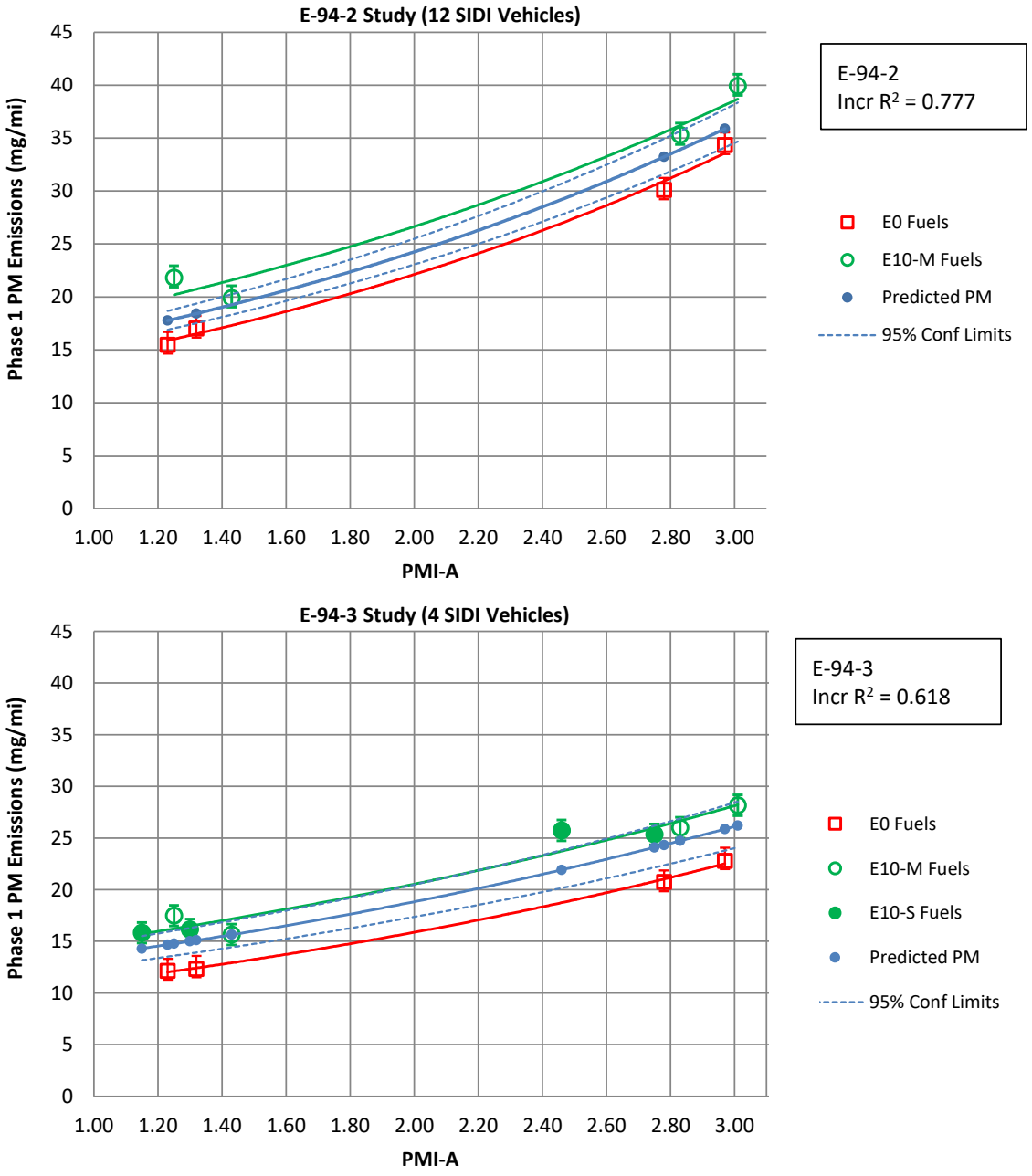


Figure B-3
Predictive Performance of PEI for E0-E10 Gasoline

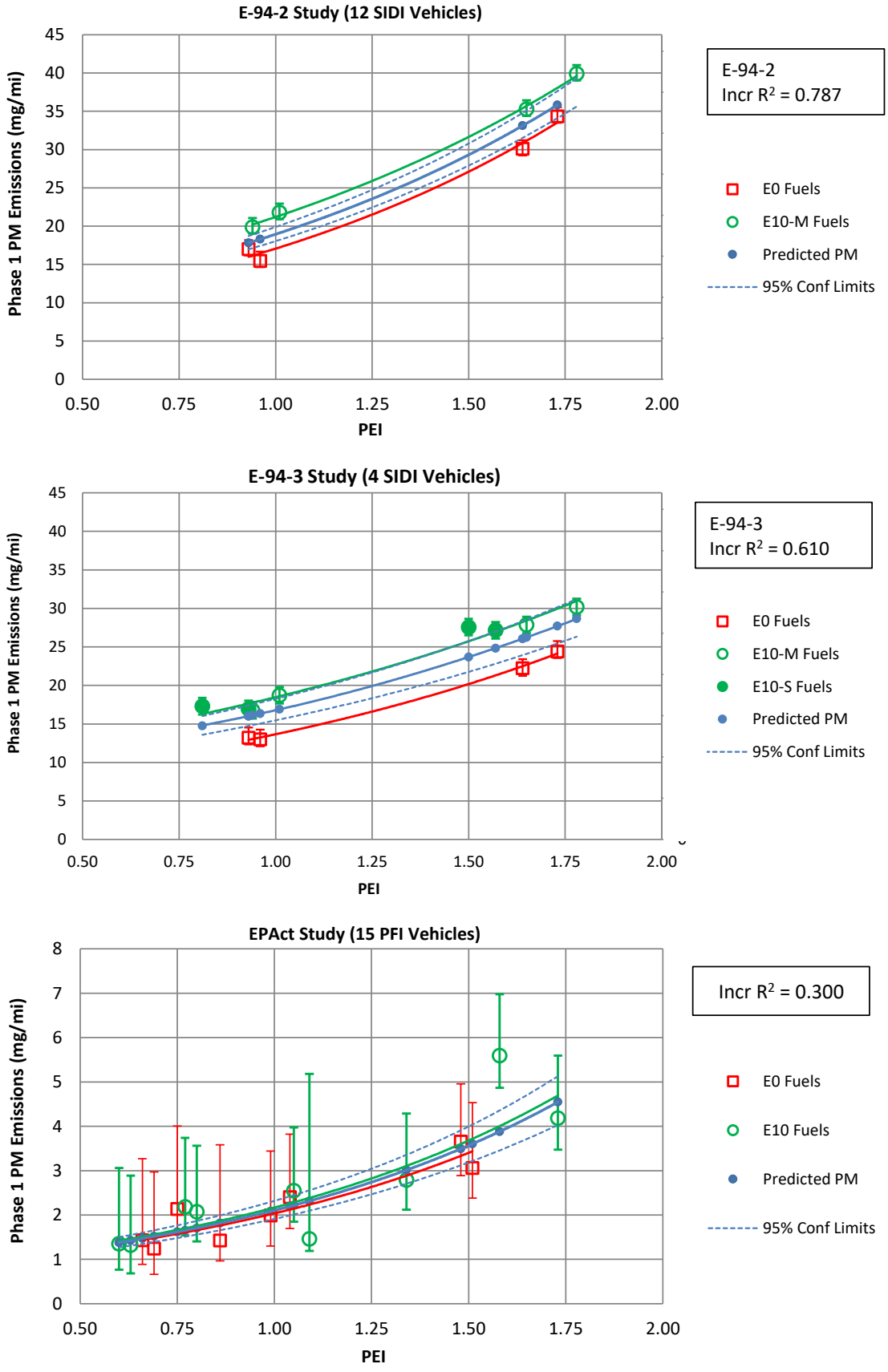


Figure B-4
Predictive Performance of 1/SP for E0-E10 Gasoline

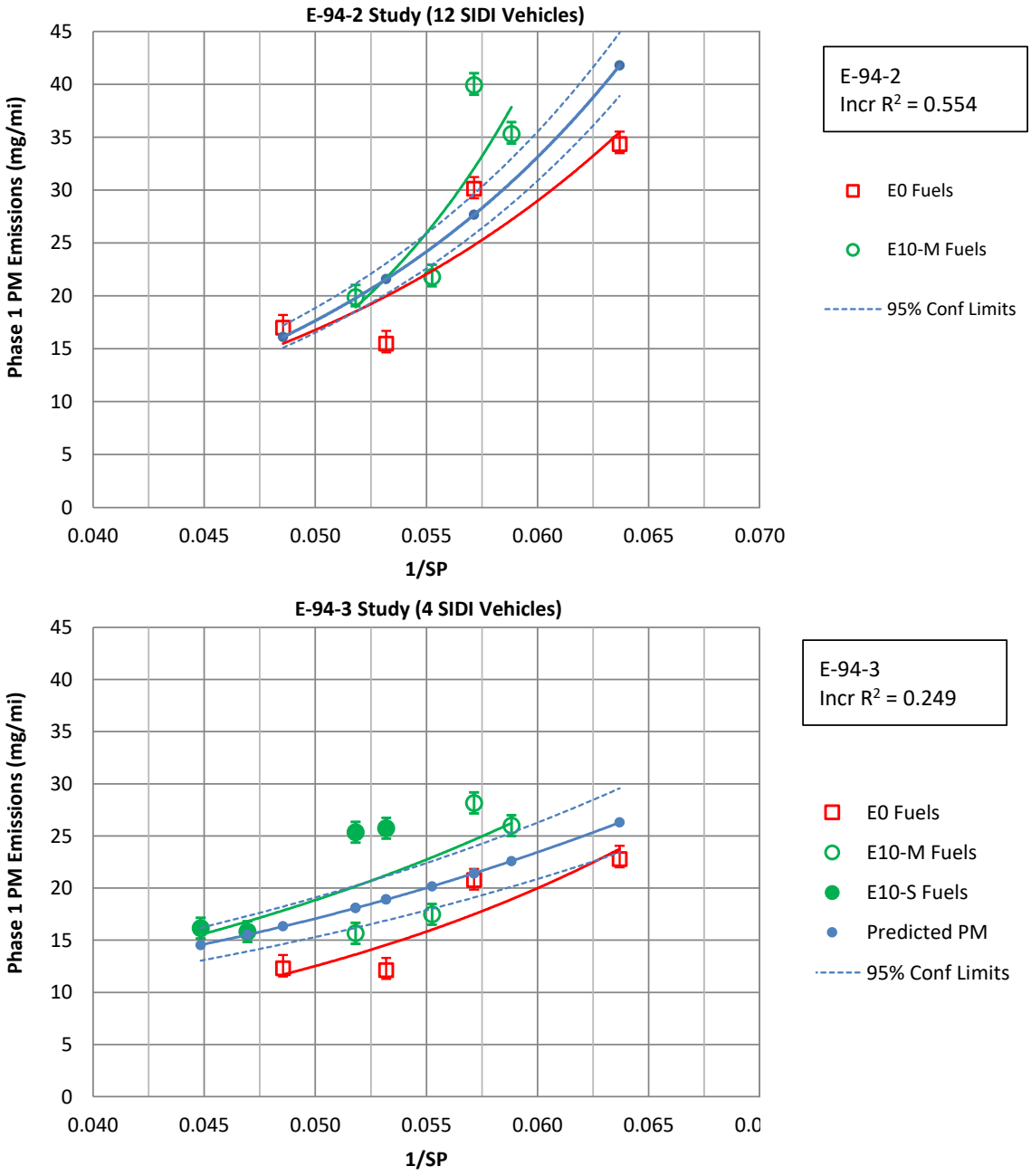


Figure B-5
Predictive Performance of OESI*-DHA MW for E0-E10 Gasoline

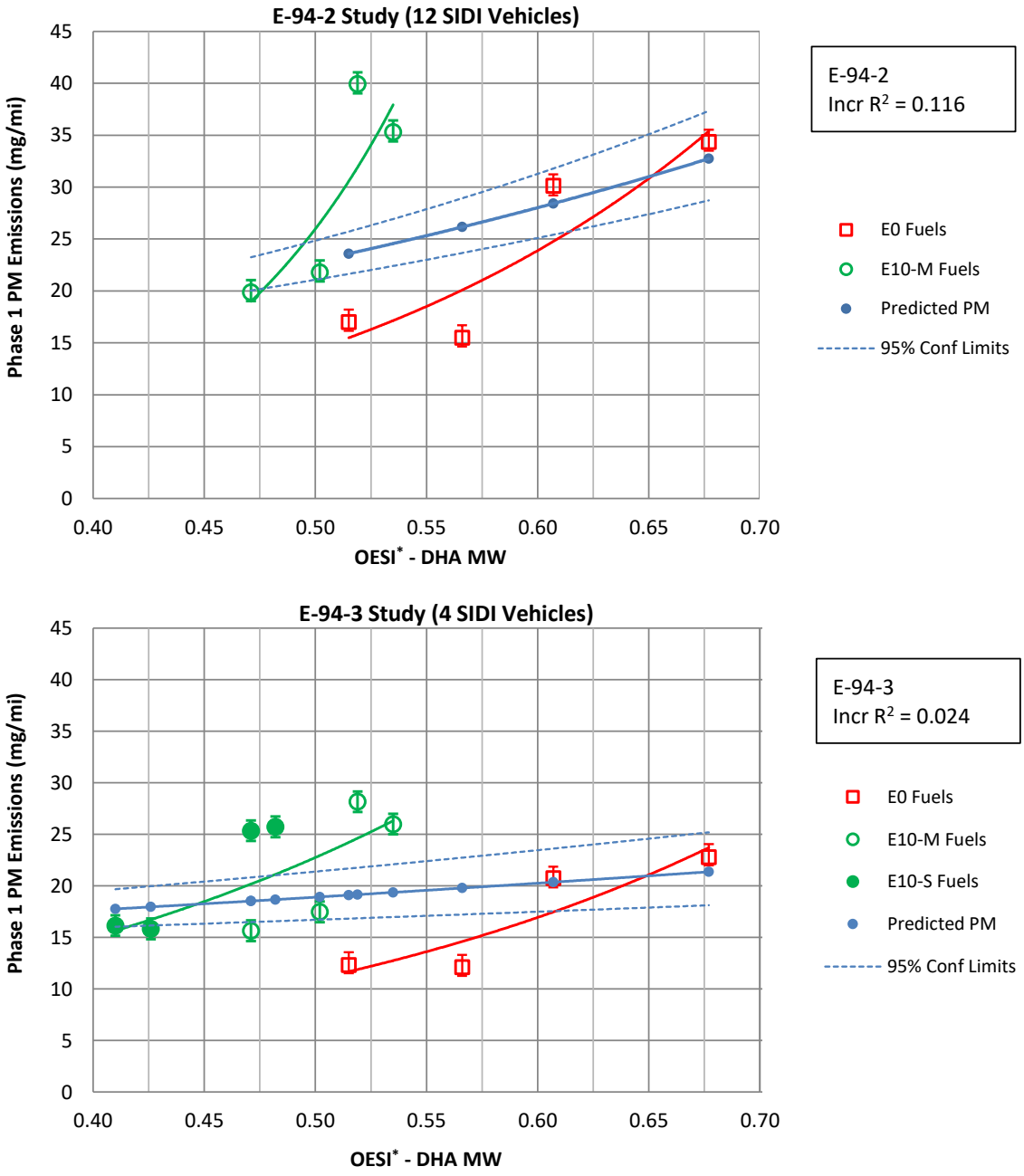


Figure B-6
Predictive Performance of PASCE for E0-E10 Gasoline

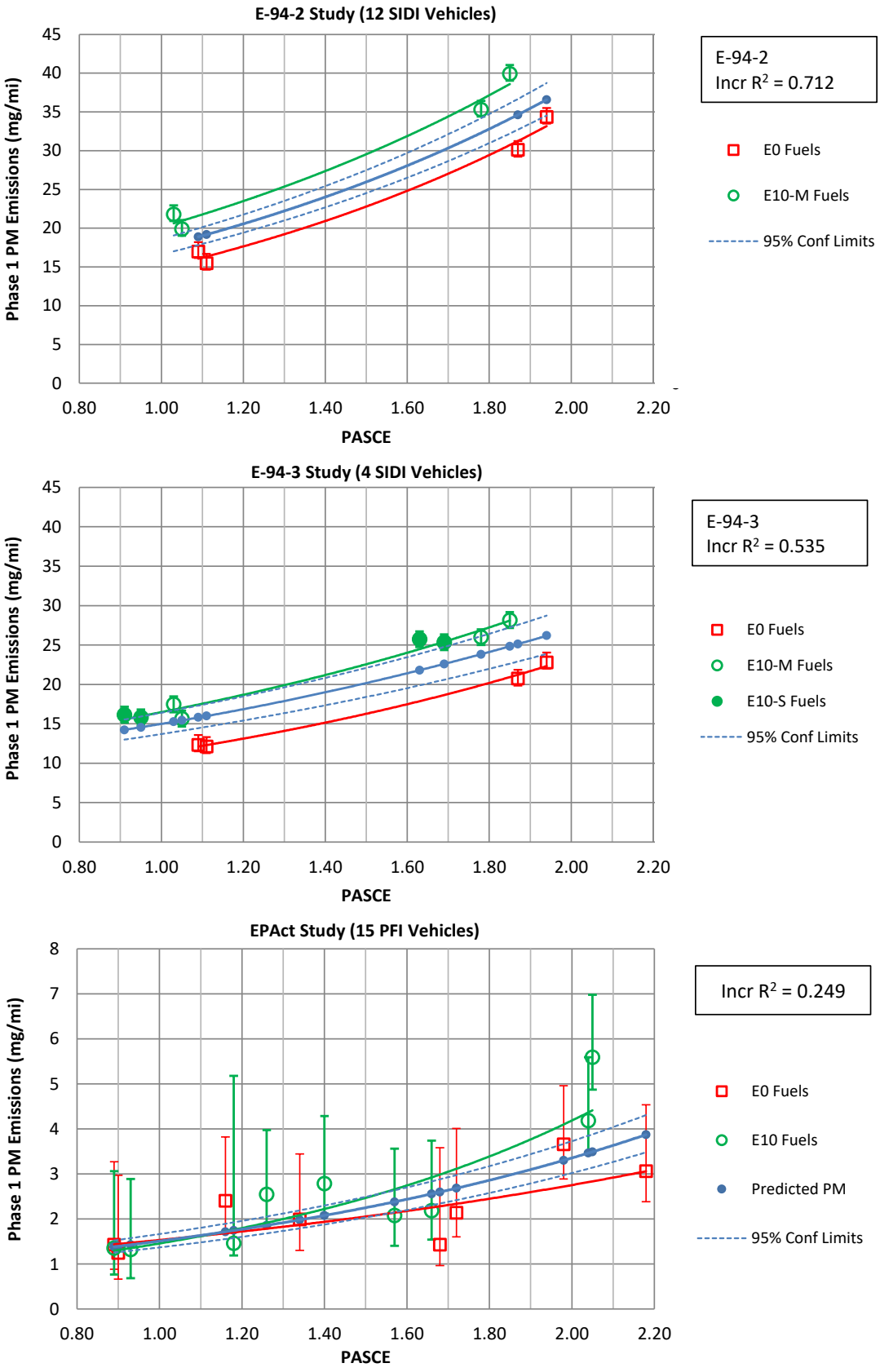


Figure B-7
Predictive Performance of E130-170 for E0-E10 Gasoline

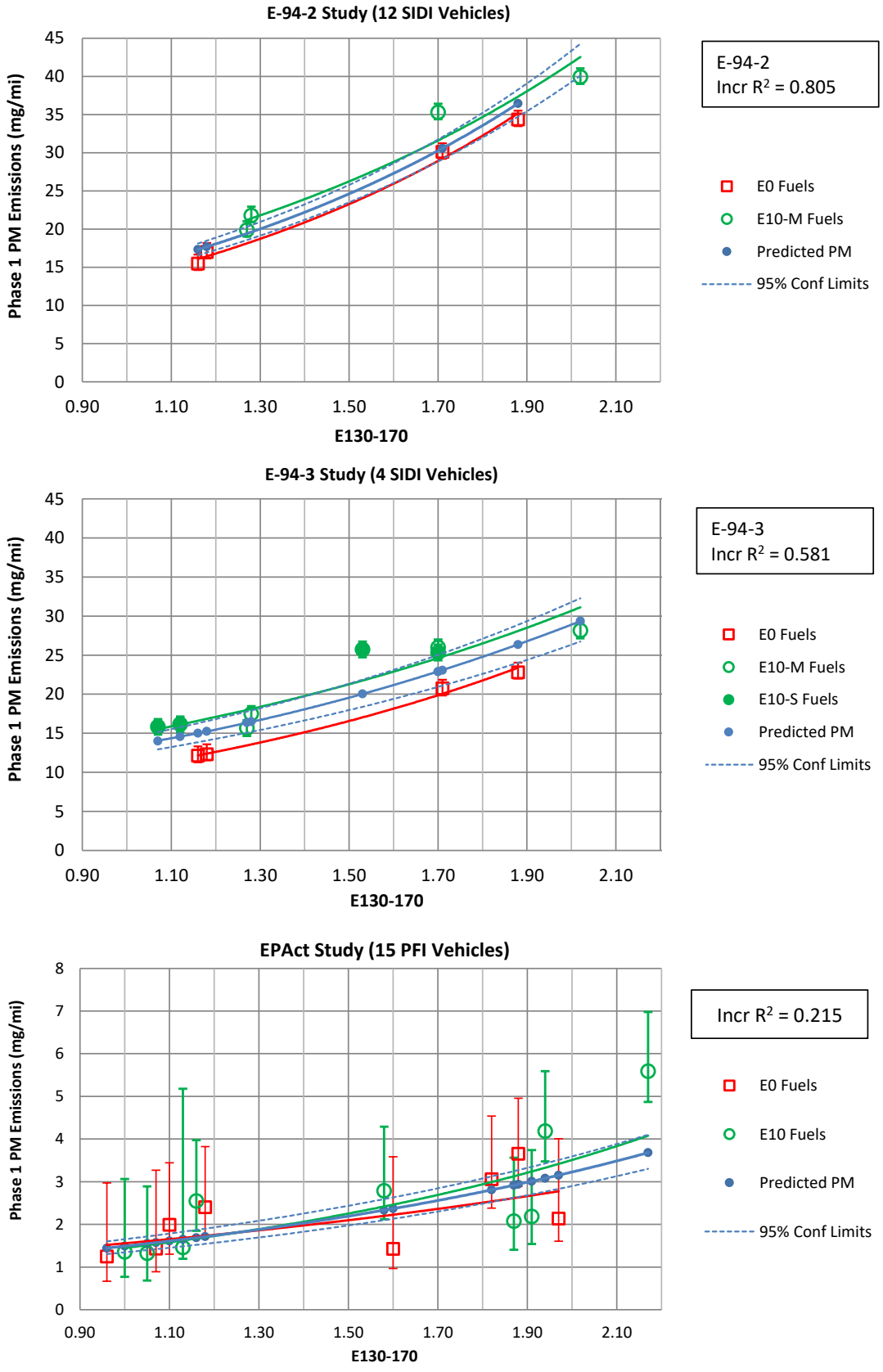
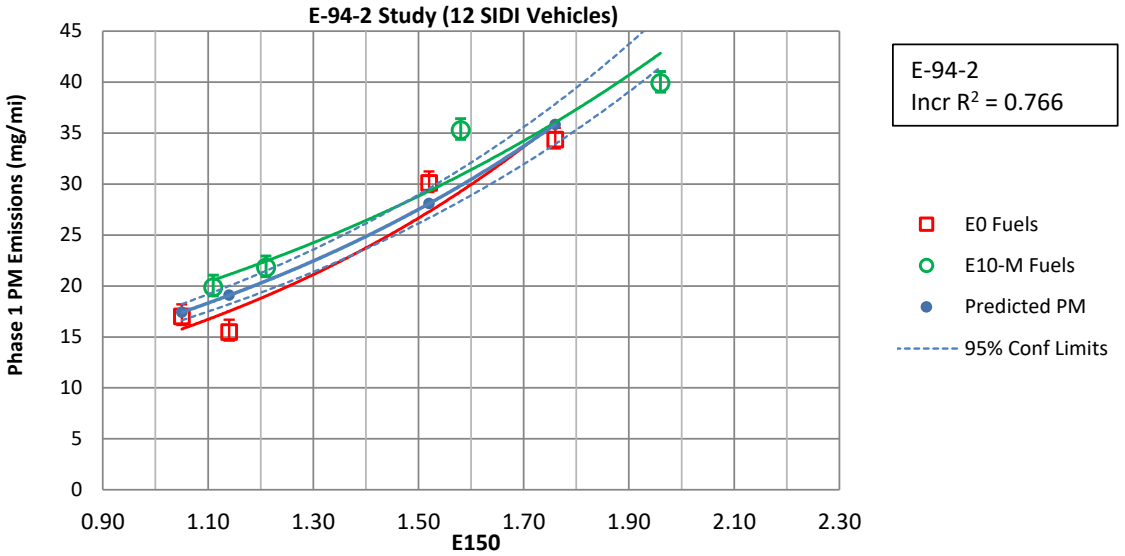
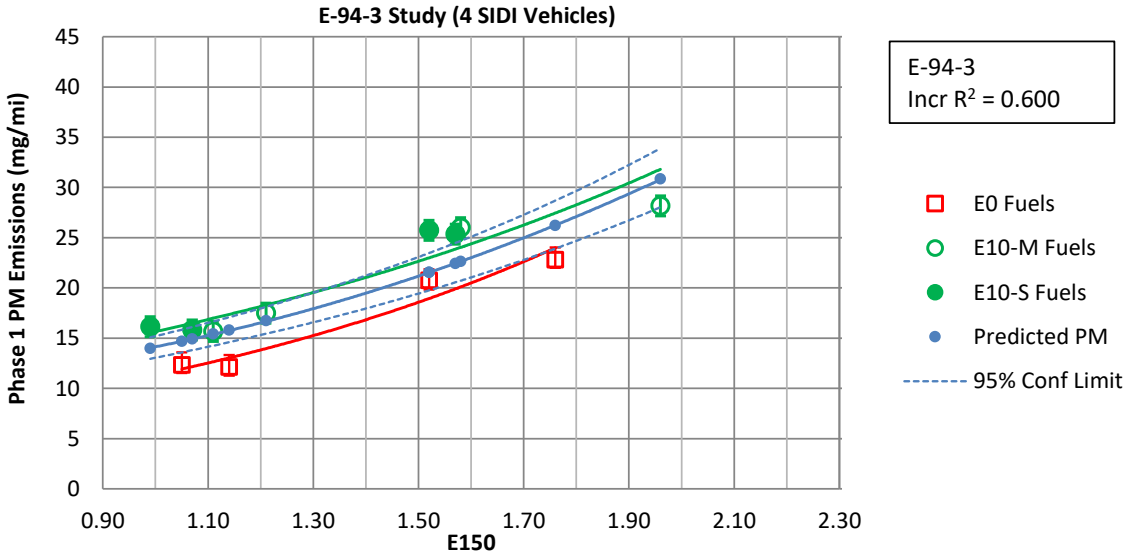
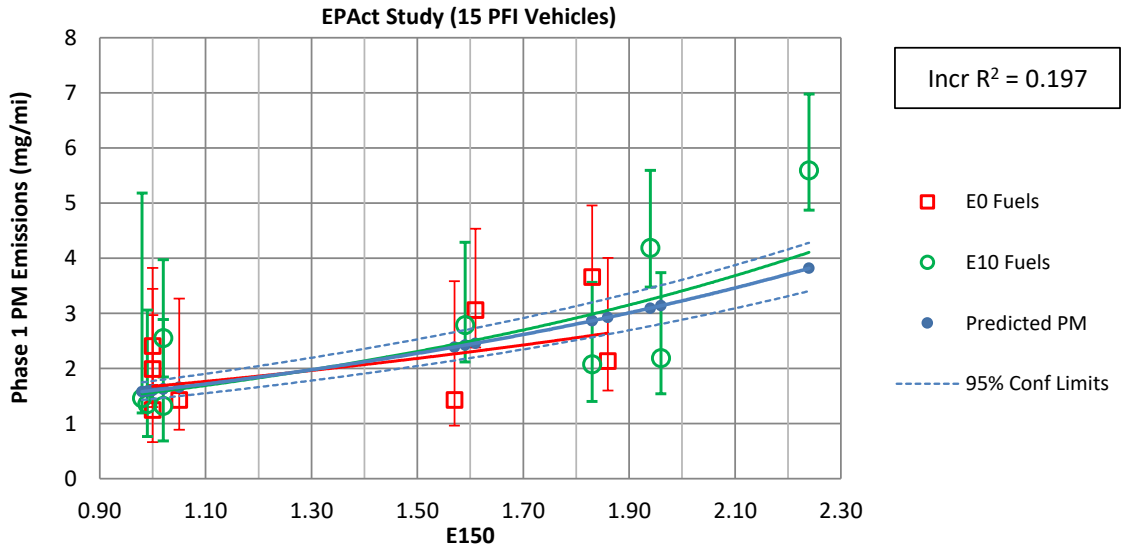


Figure B-8
Predictive Performance of E150 for E0-E10 Gasoline







APPENDIX C

PERFORMANCE RESULTS FOR E0-E20 GASOLINE

Table C-1
Summary of PMI Performance
for E0-E20 Gasoline

	EPAct
Median Phase 1 PM Emissions (mg/mi)	2.8
Incremental R^2	0.359
RMS Error	24%
mg/mi	0.6
MAE Error	60%
mg/mi	1.4
PM index of the fuel	1.83
Fuel variables adding explanatory power ($p \leq 0.01$)	
When EtOH Term is Required	T40

Figure C-1
Predictive Performance of PMI for E0-E20 Gasoline

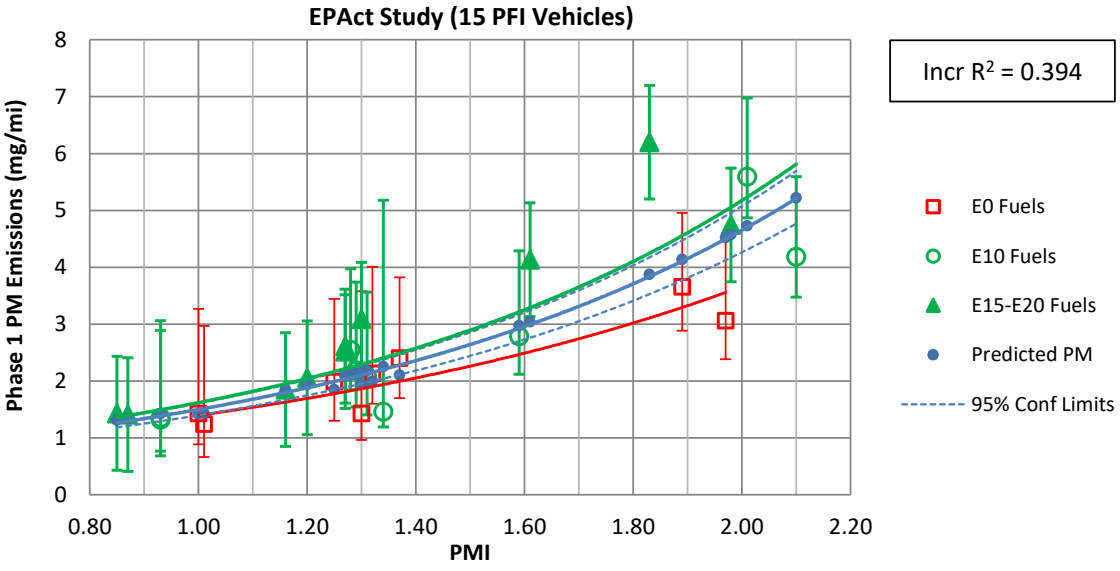


Table C-2
Summary of PEI Performance
for E0-E20 Gasoline

	EPAct
Median Phase 1 PM Emissions (mg/mi)	2.8
Incremental R^2	0.353
RMS Error	25%
mg/mi	0.6
MAE Error	74%
mg/mi	1.8
PM index of the fuel	1.09
Fuel variables adding explanatory power ($p \leq 0.01$)	
When EtOH Term is Required	T40, T70

Figure C-2
Predictive Performance of PEI for E0-E20 Gasoline

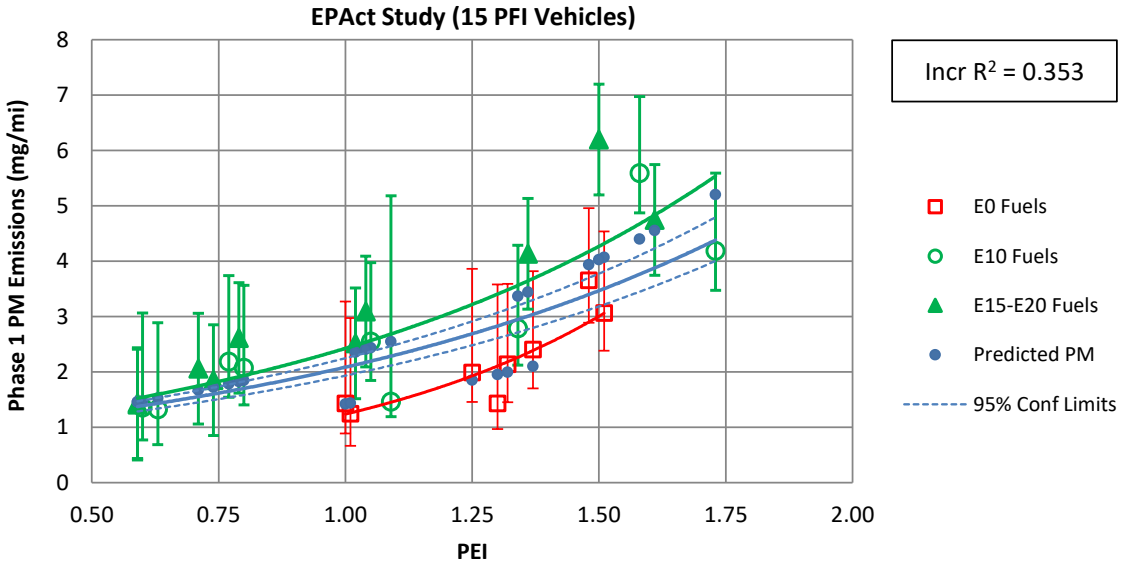


Table C-3
Summary of PASCE Performance
for E0-E20 Gasoline

	EPAct
Median Phase 1 PM Emissions (mg/mi)	0.254
Incremental R^2	36%
RMS Error	0.8
mg/mi	78%
MAE Error	1.8
mg/mi	1.42
PM index of the fuel	25%
Fuel variables adding explanatory power ($p \leq 0.01$)	
When EtOH Term is Required	Density

Figure C-3
Predictive Performance of PASCE for E0-E20 Gasoline

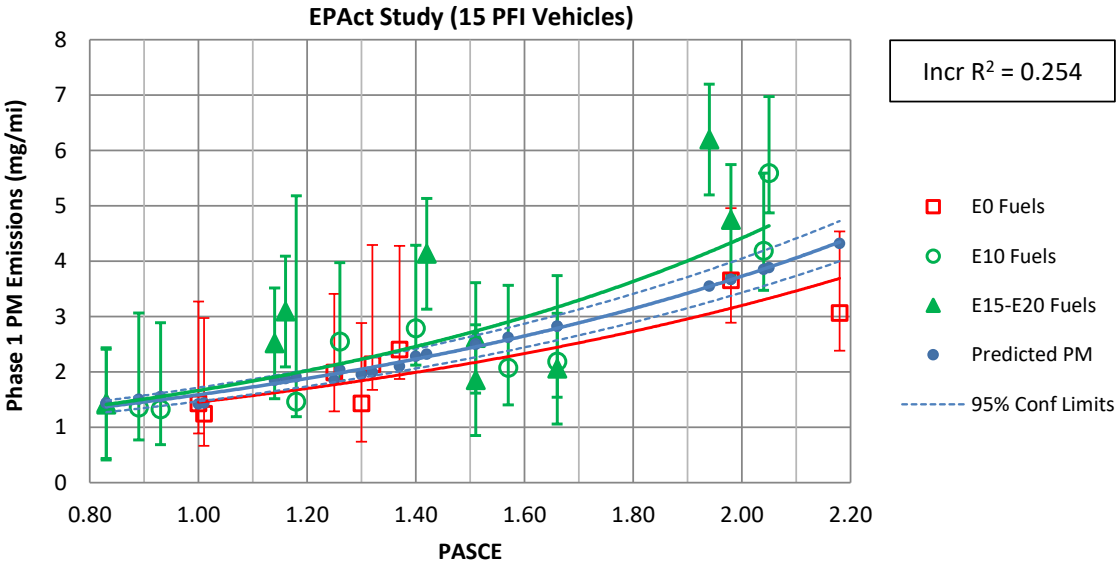


Table C-4
Summary of E130-170 Performance
for E0-E20 Gasoline

	EPAct
Median Phase 1 PM Emissions (mg/mi)	0.254
Incremental R^2	0.193
RMS Error	42%
mg/mi	1.0
MAE Error	92%
mg/mi	2.2
PM index of the fuel	1.97
Fuel variables adding explanatory power ($p \leq 0.01$)	
When EtOH Term is Required	Aromatics

Figure C-4
Predictive Performance of E130-170 for E0-E20 Gasoline

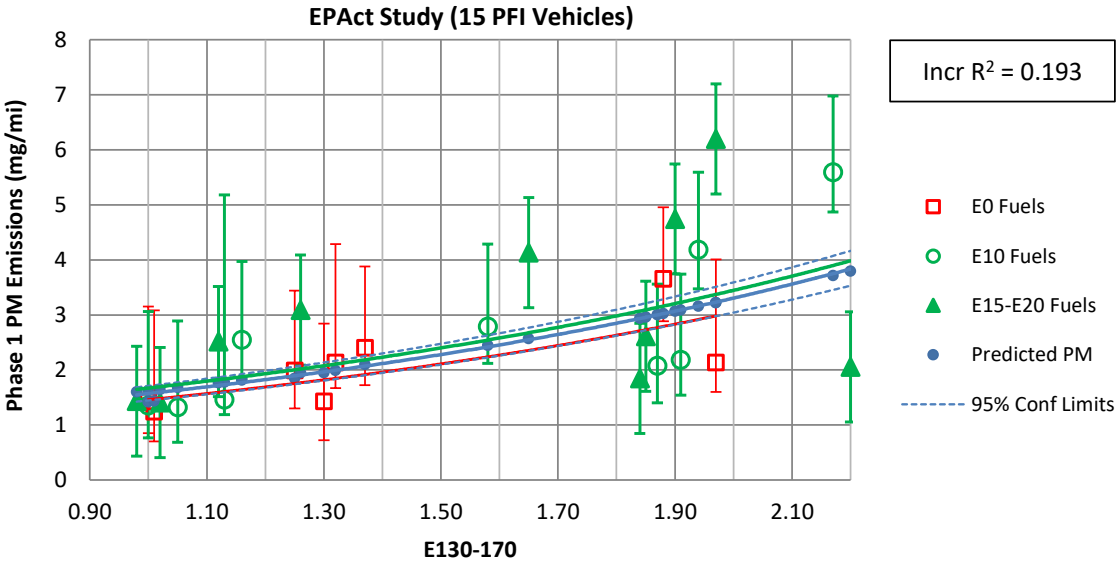
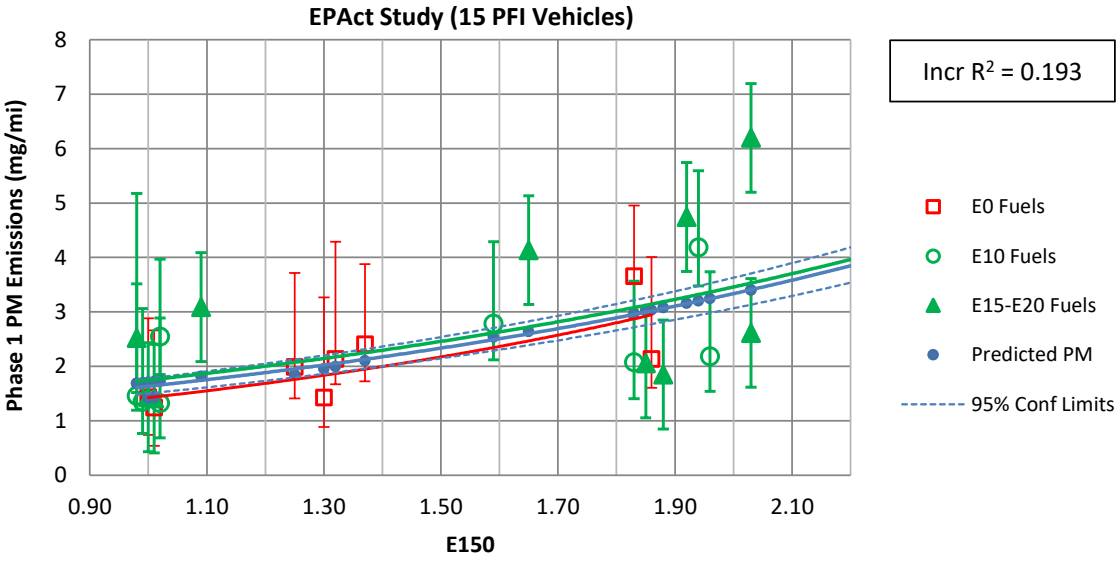


Table C-5
Summary of E150 Performance
for E0-E20 Gasoline

	EPAct
Median Phase 1 PM Emissions (mg/mi)	0.191
Incremental R^2	42%
RMS Error	1.0
mg/mi	82%
MAE Error	1.9
mg/mi	2.03
PM index of the fuel	10%
Fuel variables adding explanatory power ($p \leq 0.01$)	
When EtOH Term is Required	Aromatics, T40

Figure C-5
Predictive Performance of E150 for E0-E20 Gasoline



APPENDIX D

STATISTICAL UNCERTAINTIES IN THE PERFORMANCE MEASURES

This appendix documents the methods used to estimated statistical uncertainties for the three performance metrics that were used in the analysis to assess PM index performance.

Incremental R^2 Metric

A 95% confidence interval for the R^2 statistic can be developed from the following Eq. D-1, where SE_R is its standard error. The 95% confidence limits depend on the degree of freedom, but are approximately $R^2 \pm 2 * SE_R^2$.

$$SE_R^2 = \sqrt{4R^2(1-R^2)(N-k-1) \Big/ (N^2-1)(N+3)} \qquad \text{(Eq. D-1)}$$

The Incremental R^2 values for two PM indices are considered to statistically different when the lower 95% confidence limit for the better performing index does not overlap the upper 95% confidence limit for the poorer performing index. In the analysis, the indices were compared to the one best performing index and are described as statistically different (poorer) than the best index when the two limits do not overlap. For indices where the two limits do overlap, the observed differences are not statistically significant with 95% confidence.

This treatment is taken from Cohen et. al. (2003). *Applied multiple Regression / Correlation Analysis for the Behavioral Sciences.* p. 88. The formula is derived using a "boot strap" method, which involves repeated sampling from an assumed population to map out empirically the random variation in the R^2 statistic due to sampling fluctuations. Eq. D-1 is asymptotically valid for "large samples", which is defined in the reference as N – k – 1 > 60, where k is the degrees of freedom remaining.

It can be used for the three emissions studies where the sample size is large enough to meet the criterion or, with somewhat reduced accuracy, where the sample size approaches the criterion. As an example, the E0 fuel group for E-94-2 is not strictly large enough to meet the N – k – 1 > 60 criterion (12 vehicles * 4 fuels – 5 = 43 data points), but it is close enough for the formula to give useful, if approximate, guidance. In one case – the E0 fuels group for the E-94-3 study – the data are insufficient in number to apply Eq. D-1 and therefore statistical significance is not addressed.

RMS Metric

The 95% confidence limits for an RMS error are given by Eq. D-2, which is based on the recognition that the RMS error is the same as the standard deviation when statistical bias is not present. This is the case here because the analysis evaluates index performance in circumstances where there is no *a priori* knowledge of the presence of bias. The formula can can be applied in all subsets of the data.

$$(1\text{-}\alpha)\ CI = \sqrt{(N-1)RMS^2 \Big/ CHI^2_{\alpha/2}} \text{ to } \sqrt{(N-1)RMS^2 \Big/ CHI^2_{1-\alpha/2}} \qquad \text{Eq. D-2}$$

The formula is taken from Milton et. al. (1995). *Introduction to Probability and Statistics: Principles and Applications for Engineering and the Computing Sciences*. p. 160.

MAE Metric

Confidence limits for the MAE error are more difficult to determine for two reasons. First, the MAE calculation involves the maximum of a series of absolute values. It is likely that no closed-form solution exists for its confidence limits, although empirical confidence limits could be developed from a "boot strap" method. A diligent search did not reveal a published method.

Second, its role in the analysis is to flag the extreme error made by each index, which is used here less as a measure of comparative performance and more as a measure of risk (the worst-case error committed). The worst-case error, itself, may well be the result of systematic error (bias) in the index, rather than random sampling error. In this case, it is not amenable to the assignment of statistical limits. In the analysis, statistical uncertainties are not estimated for the MAE metric.