CRC Report No. AVFL-13c

ANALYSIS OF THE EFFECTS OF FUEL CHEMISTRY AND PROPERTIES ON HCCI ENGINE OPERATION USING A PCA REPRESENTATION OF FUELS

May 2010

COORDINATING RESEARCH COUNCIL, INC.

3650 MANSELL ROAD·SUITE 140·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, and all information is available to the public.

ANALYSIS OF THE EFFECTS OF FUEL CHEMISTRY AND PROPERTIES ON HCCI ENGINE OPERATION USING A PCA REPRESENTATION OF FUELS

Bruce G. Bunting
Oak Ridge National Laboratory
Fuels, Engines, and Emissions Research Center
and
Robert W. Crawford
Rincon Ranch Consulting

Report for CRC Project AVFL-13c

March 31, 2010

TABLE OF CONTENTS

1. PURPOSE v
2. BACKGROUND 2
3. EXPERIMENTAL FUELS 4
4. FUEL MODEL 11
5. EXPERIMENTAL ENGINE DATA 30
6. ENGINE SIMULATOR MODELS 31
7. USE OF MODELS 52
8. PARAMETRIC STUDIES 53
9. DESCRIPTION OF FUEL AND ENGINE SIMULATORS 74
10. CONCLUSIONS 82
11. RECOMMENDATIONS FOR FUTURE WORK 83
12. REFERENCES 84

FIGURES

Figure 3.1: Octane Discrepancy for Fuel B50D50 9

Figure 3.1: Distillation Curve Discrepancy for Fuel C100E20 9

Figure 4.1: Distribution of Experimental Fuels in Vector Space. 22

Figure 4.2: Predictive Power of Base Fuel Property Model. 24

Figure 4.3: Example of Paired Analysis for Ethanol Effects on Fuel Properties. 25

Figure 4.4: Primary Effects of Ethanol Blending for the Average Fuel. 26

Figure 4.5: Predictive Ability of Final Fuel Model for Base and Oxygenated Fuels. 28

Figure 6.1: Comparison of Predicted vs. Observed Values for Mode 2 Engine Simulator Models. 37-38

Figure 6.2: Ethanol Effects on Engine Performance: Experimental Data, Mode 2 (2000 rpm). 52

Figure 8.1: Coverage of Experimental and Parametric Fuels in V1, V2, and V3 Vector Space. 55

Figure 8.2: Vector Ranges for Fuels Selected to Fall Within Boundariesof Experimental Vectors and Further Selected for 87 (R+M)/2. 56

Figure 8.3: Test Phase and Control Mode Effects for Average Fuel at 2000 rpm Condition. 59

Figure 8.4: Relationship of ISFC to Score for Fuel Vector 1. 61

Figure 8.5: Comparison of ITE and ISFC for Fuel Variations, Indicating that Changes are Not Simply Due to Energy Content Changes. 62

Figure 8.6: Dependence of EVCA on Fuel Vector 2 Score 62

Figure 8.7: Dependence of EVCA on RON, MON, and Sensitivity for Parametric Fuel Set. 63

Figure 8.8: NOx Dependence on Vector 2 Value for Parametric Fuel Set. 64

Figure 8.9: dP/dCA Dependence on Vector 2 Value for Parametric Fuel Set. 65

Figure 8.10: NOx Dependence on MON, EVCA, and dP/dCA. 66

Figure 8.11: Smoke Dependence on Vector 3 Value for Parametric Fuel Set. 67

Figure 8.12: Smoke Dependence on % Aromatics,Specific Gravity, and Distillation Temperatures. 68

Figure 8.13: Vector Ranges and Fuels which Provide Desirable Operating Characteristics. 69

Figure 8.14: Tradeoffs of Olefins and Aromatics for 87 Octane (R+M)/2 Fuels of Different Sensitivity Levels. 70

Figure 8.15: EVCA vs. Sensitivity for (R+M)/2 = 87. 71

Figure 8.16: Relationship of Smoke and NOx to Aromatics for (R+M)/2 = 87. 72

Figure 8.17: Fractional Loss of ISFC and Energy Content with Ethanol Blending at 3000 RPM. 74

Figure 9.1: Graphical Output of the Fuel Panel (example for average fuel) 76

Figure A-1: Logistic Transformation of Chemistry Variables. 86

TABLES

Table 3.1: Characteristics of Selected Hydrocarbon Fuels for Phase 1. 6
Table 3.2: Characteristics of Selected Hydrocarbon Fuels for Phase 2. 6
Table 3.3: Experimental Fuels (Base and Oxygenated) Used in Engine Testing by Phase 6
Table 3.4: Correlations of Properties and Chemistries for Four Blend Components. 7
Table 4.1: Fuel Chemistry & Property Variables in Fuel Model (Independent & Dependent Variables) . 15
Table 4.2: Partitioning of Fuels Variation by Principal Components 17
Table 4.3. Definition of Vectors Forming Fuel Model. 19
Table 5.1: Definition of Engine Operating Modes 31
Table 6.1. Summary of Engine Performance Metrics 33
Table 6.2: Summary of R^2 Statistic for Mode 1 (1000 RPM) Engine Simulator Model. 41
Table 6.3: Summary of R^2 Statistic for Mode 2 (2000 RPM) Engine Simulator Model. 42
Table 6.4: Summary of R^2 Statistic for Mode 3 (3000 RPM) Engine Simulator Model. 43
Table 6.5: Summary of Oxygenated Fuel Design. 44
Table 6.6: Summary of Engine Response Model for Oxygenated Fuels: Mode 1 (1000 rpm) 48
Table 6.7: Summary of Engine Response Model for Oxygenated Fuels: Mode 2 (2000 rpm) 49
Table 6.8: Summary of Engine Response Model for Oxygenated Fuels: Mode 3 (3000 rpm) 50
Table 8.1: Average Fuel (Vector Scores = 0.0). 54
Table 8.2: Parametric Fuels Selected for Constant (R+M)/2. 55
Table 8.3: Directional Changes in Engine Performance Resulting from Control Mode Changes. 57
Table 8.4: Engine Performance Variables that are Affected by Fuel Characteristics 60
Table 8.5: Correlations of Engine Response to Fuel Parameters for (R+M)/2 = 87 at 2000 rpm 71
Table 8.6: Correlation of Engine Response Variables to % Ethanol in Fuel Blends. 73
Table 9.1: Experimental Fuel Range Covered by AVFL-13 Program. 79
Table A-1: Formulas and Coefficients Used to Compute Scaled and Transformed Values for Chemistry Variables Used in the Fuel Model 87
Table A-2: Worked Example for Fuel A100. 88
Table A-3: Normalization to Mean 0 and Standard Deviation 1, with Worked Example for Fuel A100 .. 89
Table A-4: Computation of Vector 1 Score for Fuel A100. 90
Table A-5: Computation of Vector 2 Score for Fuel A100. 90
Table A-6: Computation of Vector 3 Score for Fuel A100. 91
Table A-7: Computation of Vector 4 Score for Fuel A100. 91

ACRONYMS

ATDC..... After Top Dead Center

AVFL ..... Advanced Vehicle/Fuel/Lubricants Committee

CRC ..... Coordinating Research Council

EVCA..... Exhaust Valve Closing Angle

HCCI..... Homogenous Charge Compression Ignition

IMEP..... Indicated Mean Effective Pressure

ISFC ..... Indicated Specific Fuel Consumption

ISHC ..... Indicated Specific Hydrocarbon

$ISNO_x$..... Indicated Specific Oxides of Nitrogen

ITE ..... Indicated Thermal Efficiency

MFB50 ..... Crank angle at which Mass Fraction Burned = 50%

MON ..... Motor Octane Number

MRPR ..... Maximum Rate of Pressure Rise

NO_x ..... Nitrogen Oxides

ORNL..... Oak Ridge National Laboratory

PCA..... Principal Components Analysis

RBEI ..... Re-Breathing Early Injection

RCEI ..... Recompression Early Injection

RCSI..... Recompression Split Injection

RON ..... Research Octane Number

RVP..... Reid Vapor Pressure

SAE..... Society of Automotive Engineers

1. PURPOSE

The purpose of the research described in this report was to analyze an existing set of experimental engine and fuel data which had been contracted by the Advanced Vehicle/Fuel/Lubricants (AVFL) Committee of the Coordinating Research Council (CRC) to study gasoline range fuel effects on Homogenous Charge Compression Ignition (HCCI) combustion. The analysis was aimed at developing broad, overall models for the engine response to changes in engine control and fuel variables, to exercise the models in a series of parametric studies, to structure the models in such a way that they could be used by others for additional parametric studies, and to make recommendations for future research in fuels and advanced combustion engines, based on the outcome of the analysis.

This analysis differed from a previous analysis (1, 2) in that in the current work an emphasis was placed on overall models and that Principal Components Analysis (PCA) was used to represent the fuels. It is well-known that the physical properties of petroleum fuels are frequently correlated to each other through the underlying chemistry that connects them and that these correlations often complicate or confound studies of fuels. This analysis uses PCA as a solution to this problem. PCA creates new vector variables that carry the correlations present in the fuels through subsequent analysis, without requiring the artificial choice of one (or a few) of the correlated variables to act as surrogates. In this approach, all of the original variables contribute to the analysis through their role in defining the vectors themselves, thereby retaining the highest fidelity to the original data. Further, the vector variables are statistically efficient. In circumstances where correlations are present among the original variables, there are fewer degrees of freedom (free choices) actually present in the data than the number of variables would suggest, and a relatively smaller number of vectors may be needed to represent the data.

In this analysis, experimental variation was accounted for, as much as possible, by including variables related to variation in experimental conditions, including set-points, environmental factors, phase of research, and control methodology of the engine. This allowed the construction of as few models as possible, although separate models for each operating point were needed.

The results presented in this report capture the major trends (and limitations) of the data and also can be considered as a guide for the further use of the supplied models for other studies. The authors are also available for consultation if needed in applying these models to other studies. This may be important since it is fairly easy to overstep the experimental data when using the models.

2. BACKGROUND

The following is a brief background of the history of this project as a reminder of the fuels and engine conditions selected and how it evolved over time. Further details can be obtained from two CRC reports (1, 2) and one SAE paper (3). From 2006 through 2008, CRC sponsored a series of engine tests in order to help understand the effects of fuel chemistry and properties on

HCCI engine performance. The overall project was contracted to AVL by CRC, which, in turn, subcontracted with Battelle Memorial Institute for statistical and experimental design support and with Oak Ridge National Laboratory (ORNL) for a literature review and recommendations of fuel properties and chemistry of interest. The development of the experimental plan was a collaborative effort between CRC AVFL, AVL, Battelle, and ORNL.

It was recommended that the fuel matrix cover the variable space defined by blends of the four main gasoline blending streams (reformate, alkylate, cat cracker gasoline, and straight run gasoline) (5). Originally, the recommended independent variables for the fuels were RON (Research Octane Number), MON (Motor Octane Number), (RON+MON)/2, sensitivity (RON-MON), % aromatic, % olefin, % iso-paraffin, % normal paraffin, and % cyclo-paraffin. Three RON values of about 92, 80, and 70 were chosen, with as wide a range of MON, sensitivity, and chemistry as could be achieved by combining the four blend streams. A total of ten fuels was selected from a series of 56 possible fuels blended by ConocoPhillips.

Volatility, as defined by boiling point distribution or RVP (Reid Vapor Pressure), was not expected to play a large role for gasoline range fuels operating under the steady state, warm engine conditions selected for the engine testing, and it was recommended that the four blending streams be matched as closely as possible for boiling points in order to eliminate this from consideration. However, the selected blending streams did differ somewhat in boiling points. The final fuels were also doped with n-butane to control RVP to about 7 psi for vapor safety considerations.

In the second phase of the project, ethanol was splash-blended into some of the hydrocarbon fuels, which were selected based on CRC group interest augmented with some statistical guidance. In this, the physical properties of the oxygenated fuel blends were allowed to vary without constraint. Octanes and RVP were allowed to increase, and boiling points were allowed to decrease, compared to the characteristics of the base hydrocarbon fuels. In this sense, the Phase 2 oxygenated fuels form a separate set of fuels from their hydrocarbon counterparts. The engine tests for these were run in two Phases (2a and 2b) with a major engine rebuild occurring between them.

Three operating conditions for the engine were selected to be 1000 rpm at 1.5 bar IMEP, 2000 rpm at 3.0 bar IMEP, and 3000 rpm at 5.5 bar/deg pressure rise rate. Three control modes for the engine included: port fuel injection with negative valve overlap (known as recompression early injection - RCEI); port fuel injection with a second exhaust valve event (known as re-breathing early injection - RBEI), and direct injection with a second injection during the negative valve overlap period (known as recompression split injection - RCSI). In all cases, the crank angle at which the Mass Fraction Burned is equal to 50% (MFB50) was controlled to 5° ATDC (After Top Dead Center). After analysis of the data from the first ten fuels was completed by AVL and Battelle, it was expected to continue the research by blending additional fuels to cover ranges found to be of particular interest. An indolene reference fuel was also periodically run in order to set baseline engine control parameters and to track engine condition and other experimental

variation. Two additional series of tests were run (designated Phase 2a and 2b), primarily to study ethanol blending effects. Battelle performed statistical analysis of the data, mainly using two factor fuel effect analyses by engine operating mode and engine control mode, while also uncovering overall trends which were present in the data. After completion of the Phase 1 report (1), the experimental plan was revised to include effects of ethanol blending in the Phase 2 testing. The experimental data have several characteristics that make analysis difficult. First, the data were taken in three separate series (Phase 1, Phase 2a, and Phase 2b), which took place during 2006, 2007, and 2008 respectively. During this time, the engine underwent minor maintenance, one major rebuild, a change of project engineer, and normal variations of engine wear condition and in weather. Second, the experimental design was altered to include ethanol effects in Phases 2a and 2b, rather than additional study of petroleum chemistry effects. Third, the experimental plan was altered to eliminate one of the control modes for later test phases. Finally, there may have been some shift in the fuel properties or chemistry that were of most interest for the analysis. The end result of these changes was an unbalanced experimental design that makes it difficult to compare fuel variables across the entire range of data.

Despite these challenges, the authors of this report felt that it would be feasible to model the data in its entirety using a different statistical approach in order to identify large, overall fuel and engine trends to help guide future research in this area. The CRC AVFL Committee agreed to fund this work under Project AVFL-13c, the results of which are presented in this report. This analysis was based on applying PCA to represent the fuels, on using a series of dummy variables to represent sources of variation such as test phase, engine operating mode, and engine control mode, and by including control target variables in the model to allow correction for experimental variation. The scope of the work included delivering the models to CRC in active Excel™ format so that members can do further analysis and parametric studies as desired. This report is a summary of the modeling and PCA analysis of this data with a focus on the major trends in engine and fuels that are observed. As is true of all analytical work, the results may be imperfect because of correlations in the data that prevent full separation of the variables of interest and because not all experimental variation could be assigned to an independent variable.

3. EXPERIMENTAL FUELS

3.1 Characteristics of the Experimental Fuels

The experimental fuels used in the AVFL program were created to explore a wide range of potential fuel property and chemistry effects on HCCI engine performance. As described in the first CRC report (1), a literature review (4) suggested that gasoline engine performance would be affected by octane number – including RON, MON, and sensitivity – in addition to characteristics of the fuel's distillation curve and chemistry. Fuel oxygen content was also added to the list of potential fuel effects for Phase 2 of the program.

To test such effects, a series of 56 candidate gasoline-range fuels were identified in Phase 1 that could be created by blending four common refinery blend streams designated as A, B, C, and D.

Properties of these candidate fuels were estimated by a blending model. The blend streams were chosen to be typical of the refinery stocks from which gasoline is commercially blended:

- Stream A: a reformate having high aromatic content (about 55%)
- Stream B: an alkylate having high iso-paraffin content (about 88%)
- Stream C: a cat-cracker gasoline having high olefin content (about 45%)
- Stream D: a light straight-run gasoline (about 24% n-paraffin)

Of the 56 candidates, several fuels were also based on road (commercial) gasoline blends or were splash-blended with kerosene, but such fuels were excluded from the engine testing by the CRC committee. Following the exclusion, 37 candidate fuels remained from which the ten initial experimental fuels were chosen.

Table 3.1 lists the fuels selected for engine testing in Phase 1 along with their chemistry and properties and also includes D100, which was not tested (but used as a blendstock), and indolene (used as a reference test fuel), which was not used in construction of the models. In this table, fuels are described in terms of blending percentages – e.g., A75D25 is a fuel targeted as a 75:25 blend (by volume) of stocks A and D. The ten selected base fuels involve individual two-way substitutions between blend streams, meaning substitutions of streams A vs. D, of B vs. D, and of C vs. D. One additional base fuel was created from a three-way blend of A, C, and D for phase 2a. Of the four blend stocks, Streams A, B, and C were also tested alone (as A100, B100, and C100) and as blends with Stream D, while Stream D was too low in octane to be an experimental fuel in its own right. After blending, the varying volatility of the fuels was controlled by blending n-butane as needed to achieve a Reid Vapor Pressure (RVP) of 7 psi.

For Phase 2, oxygenated fuels were created by splash-blending ethanol into selected base fuels at volumetric percentages ranging from E10 up to E30. These fuels are listed in Table 3.2, where an oxygenated fuel comprised of 90% A75D25 (by volume) splash-blended with 10% ethanol is described as A75D25E10. Five series of oxygenated fuels were created from different base fuels and splash-blended with ethanol to create a total of ten oxygenated fuels.

While the 11 base fuels and ten oxygenated fuels were used in engine testing, two of the fuels and their engine tests have been dropped from this analysis, as discussed in Section 3.2. Fuel B50D50 was dropped because the measured T50 and T90 values were higher than those of the supposed “parent” B100 and D100 fuels and because the reported octane data appeared to be incorrect. Fuel C100E20 was dropped because the distillation curve appeared to be incorrect. These exclusions leave a dataset covering 15 base fuels, including the ten base fuels tested in Phase 1 and the five base fuels re-blended and retested in Phase 2. The dataset also covers nine oxygenated fuels created with varying ethanol contents from four base fuels. With the exclusion of C10020, only one oxygenated fuel series remains in Phase 2a testing, while three such series remain in Phase 2b testing. The sequence of fuels for each test phase is shown in Table 3.3. Some of the fuels were run several times and were re-measured, resulting in some slight

chemistry and property changes between Tables 3.1 and 3.2. Additionally, a different C100 base stock was used in Phase 2a vs. Phase 1.

Table 3.1: Characteristics of Selected Hydrocarbon Fuels for Phase 1.

Test Results	A100	B100	C100	D100	B50D50	C85D15	B77D23	B30D70	A79D21	C50D50	A33D67	Indolene
RON	95.5	89.1	91.5	52.8	58.9	87.7	81.0	64.7	84.0	78.0	65.5	95.6
MON	84.8	89.3	80.3	50.4	61.1	77.3	81.9	63.7	76.2	71.5	62.8	89.0
Sensitivity	10.7	-0.2	11.1	2.4	-2.1	10.4	-0.8	1.1	7.9	6.4	2.7	6.6
RVP (psi)	7.3	7.0	7.5	6.5	7.5	7.2	7.1	7.3	6.9	7.3	7.0	9.1
Density (60F/60F)	0.796	0.698	0.717	0.757	0.752	0.734	0.713	0.738	0.781	0.741	0.766	0.743
T10, F	175	154	134	190	187	139	159	173	182	147	191	127
T50, F	266	212	179	276	321	213	224	254	266	242	272	221
T90, F	327	338	256	345	421	297	337	344	330	320	339	322
% Butane	2%	4%	0%	7%	8%	1%	7%	6%	4%	4%	6%	1%
% n-Paraffins w/o C4	10%	1%	4%	24%	21%	6%	6%	16%	14%	12%	20%	7%
% i-Paraffins	27%	88%	30%	27%	42%	29%	77%	51%	31%	31%	29%	57%
% Cyclo-Par	3%	3%	10%	23%	20%	11%	7%	15%	8%	16%	17%	6%
Olefins	2%	1%	45%	2%	1%	37%	1%	1%	2%	18%	2%	1%
Aromatics	55%	3%	11%	16%	8%	17%	3%	11%	41%	18%	26%	27%

Table 3.2: Characteristics of Selected Hydrocarbon Fuels for Phase 2.

Test Results	C100	C100E20	A79D21	A79D21E10	A79D21E20	A79D21E30	A50C20D30	C50D50	C50D50E15	C50D50E30	B100	B100E15	B100E30	B77D23	B77D23E15	B77D23E30
RON	86.0	95.1	83.4	88.5	94.7	99.1	82.4	77.2	87.3	94.6	90.0	100.4	105.9	82	94.2	101.2
MON	76.6	82.4	76.5	80.4	83.4	86.9	75.8	72.6	78.8	82.5	89.9	92.7	94	81.7	88.5	90.8
Sensitivity	9.4	12.7	6.9	8.1	11.3	12.2	6.6	4.7	8.5	12.1	0.0	7.7	11.9	0.3	5.7	10.4
RVP (psi)	6.8	7.9	6.6	7.5	7.2		6.8	7.3	8.2	7.6	7.4	8.1	7.6	7.4	8.1	7.7
Density (60F/60F)	0.713	0.727	0.781	0.781	0.782	0.782	0.768	0.741	0.749	0.756	0.698	0.712	0.726	0.713	0.724	0.734
T10, F	140	129	188	153	156	158	166	148	137	143	159	141	145	167	145	148
T50, F	175	149	267	261	252	171	255	245	192	168	216	165	162	225	215	165
T90, F	240	233	332	327	323	319	325	323	317	311	354	316	279	343	335	324
% Butane	0%	0%	4%	3%	3%	2%	4%	4%	3%	3%	5%	4%	3%	7%	6%	5%
% n-Paraffins w/o C4	8%	7%	14%	12%	11%	9%	14%	13%	11%	9%	0%	0%	1%	6%	5%	4%
% i-Paraffins	33%	26%	32%	29%	25%	21%	31%	31%	27%	22%	89%	75%	63%	74%	62%	52%
% Cyclo-Par	11%	9%	8%	7%	6%	5%	10%	15%	13%	10%	1%	1%	1%	6%	5%	4%
% Olefins	36%	28%	2%	1%	1%	1%	7%	17%	15%	12%	0%	0%	0%	1%	1%	1%
% Aromatics	10%	9%	41%	37%	33%	28%	35%	20%	16%	14%	5%	4%	3%	8%	7%	6%
% Ethanol	0%	21%	0%	11%	22%	35%	0%	0%	16%	30%	0%	16%	30%	0%	16%	29%

Table 3.3: Experimental Fuels (Base and Oxygenated) Used in Engine Testing by Phase.

	TEST PHASE				
base fuel	0% EtOH	10% EtOH	15% EtOH	20% EtOH	30% EtOH
A100	1				
A79D21	1, 2A	2A		2A	2A
A50C20D30	2A				
A33D67	1				
B100	1, 2B		2B		2B
B77D23	1, 2B		2B		2B
B50D50	1				
B30D70	1				
C100	1, 2A			2A	
C85D15	1				
C50D50	1, 2B		2B		2B

In any set of fuels, there are usually a large number of correlations that exist between the variables to potentially confound subsequent analysis, and this fuel set is no exception. It is

perhaps simplest to consider correlations among the four blending streams, A100 through D100, as shown in Table 3.4.

Table 3.4: Correlations of Properties and Chemistries for Four Blend Components.

	RON	MON	Sens	RVP	Density	T10	T50	T90	% nC4	% n-P w/o C4	% i-P	% Cyclo-P	Olefins	Aromatics
RON	1.00													
MON	0.96	1.00												
Sens	0.51	0.24	1.00											
RVP	0.89	0.74	0.80	1.00										
Density	-0.10	-0.26	0.46	-0.03	1.00									
T10	-0.67	-0.65	-0.32	-0.75	0.68	1.00								
T50	-0.54	-0.55	-0.19	-0.62	0.78	0.98	1.00							
T90	-0.45	-0.28	-0.70	-0.77	0.29	0.81	0.78	1.00						
% nC4	-0.84	-0.69	-0.80	-0.99	0.09	0.79	0.68	0.84	1.00					
% n-P w/o C4	-0.87	-0.93	-0.12	-0.70	0.58	0.84	0.79	0.43	0.68	1.00				
% i-P	0.25	0.51	-0.70	-0.14	-0.70	-0.30	-0.35	0.31	0.16	-0.61	1.00			
% Cyclo-P	-0.94	-0.98	-0.23	-0.68	0.10	0.49	0.37	0.13	0.61	0.85	-0.47	1.00		
Olefins	0.30	0.14	0.61	0.65	-0.36	-0.78	-0.77	-0.98	-0.74	-0.34	-0.32	0.02	1.00	
Aromatics	0.28	0.12	0.59	0.29	0.93	0.42	0.57	0.15	-0.20	0.22	-0.55	-0.28	-0.28	1.00

These correlations describe several characteristics in this set of blending streams, but one also has to interpret or classify these characteristics to understand their general application to gasoline blending streams versus their limited application to this particular sampling of blend streams. The investigators' attempt to do this is shown below.

- RON and MON are highly correlated, probably because of the wide range of RON and MON resulting from inclusion of the low octane D100 blendstock in these correlations.
- RON goes down with n-paraffins content (n-P) and cycloparaffins content (cyclo-P). This is universally true for these blending streams, because blend stream D (straight run) was the only stream which significantly lowered octane and contained the largest fractions of n-P and cyclo-P.
- RON not highly correlated to isoparaffins (i-P), olefins, nor aromatics content. This shows that the blend streams were capable of balancing between these three choices for building octane.
- Sensitivity goes down with i-P and up with olefins and aromatics while cycloparaffins (cyclo-P) have little effect. This is universally true for gasoline fuels.
- T10, T50, and T90 are strongly correlated to each other. This is probably always true for refinery components with broad boiling point distributions. It would be less true if pure components, such as toluene, were included in the fuels.
- The olefin stream had lower boiling points and the straight run had higher T10 and T50. The authors are not sure if this is typical for these streams or just for this representation of these streams, but it must be kept in mind when interpreting the results. This is an important consideration, since variation will be assigned to all variables that change between blending streams, including boiling points. It should also be kept in mind that there is a general trend for pure hydrocarbons that octane values decrease with higher boiling points.

Overall, the blend streams and fuels selected provided the ability to study octane, chemistry, and sensitivity effects in two component blends and two component blends plus ethanol. The fuels and blendstreams did not include the study of three component hydrocarbon blends (although one three component fuel was tested – A50C20D30) and did not allow the full separation of chemistry and boiling points.

3.2 Fuels Excluded from the Analysis

Two fuels, B50D50 and C100E20, appear to have been inconsistently or incompletely characterized with respect to chemistry and properties. These fuels and the related engine test data were therefore excluded from the analysis in their entirety. Note that these fuels were retained in the original analysis since they were not examined for consistency of their properties and chemistry.

Fuel B50D50 was found to be an outlier with respect to its T50, T90, and octanes when compared to the other fuels and to an initial fuel model that otherwise fit the fuel dataset well. When the fuels are plotted against the percentage of stream D in the blend (see Figure 3.1), one sees that D (consisting of mixed paraffins) was the main blending stream for adjusting octanes in the fuels. The RON and MON octane values for the experimental fuels follow a well-defined trend with percent D, except for fuel B50D50 which falls well below the trend. Similar discrepancies were noted for sensitivity and the T90 boiling point. Taken together, this evidence indicates that the characterization of B50D50 may be in error or, more likely, that it is not a blend of B100 and D100. The discrepancies were substantial and had a large effect on the parameters of the fuel model. Therefore, the fuel was excluded from the analysis presented here.

In the subsequent analysis of the oxygenated fuels, a discrepancy in the distillation curve was discovered for fuel C100E20, which was also excluded from the analysis. As Figure 3.2 shows, the C100E20 distillation curve is depressed compared to that for C100 until more than 80 percent of the fuel has been distilled. If C100E20 were a 20 percent splash blend of ethanol with C100, the distillation curve should return to the C100 values much sooner than this. Thus, the C100E20 fuel must be something else, perhaps a 20 percent blend with another base fuel or a higher ethanol blend with C100, or its data may be subject to another error. Because the discrepancies preclude determining the base fuel and ethanol percentage with certainty, the fuel and its related test data were excluded from the analysis.

Figure 3.1: Octane Discrepancy for Fuel B50D50.

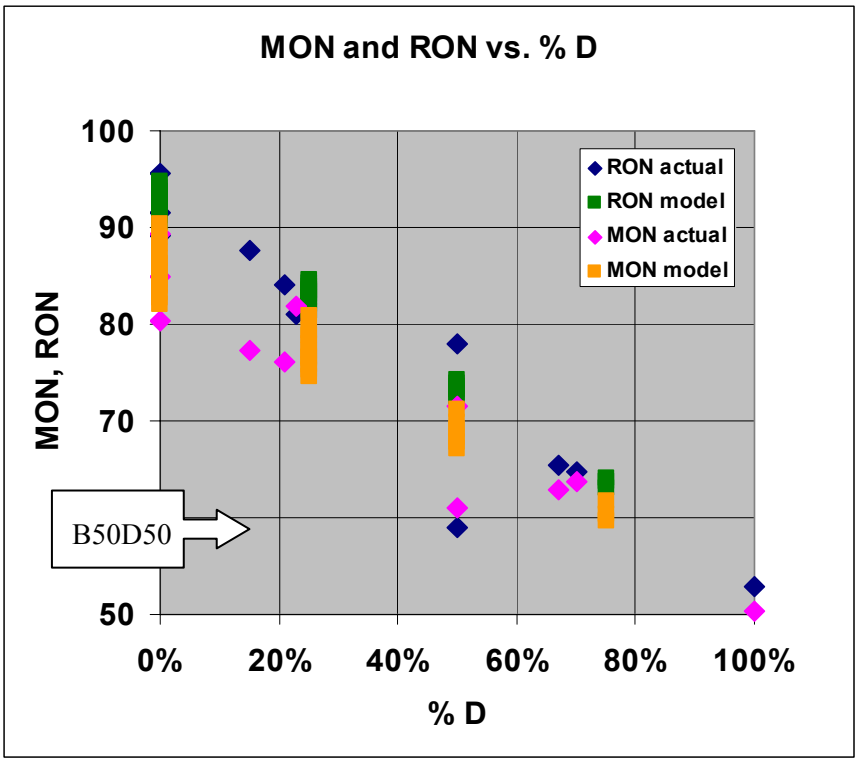
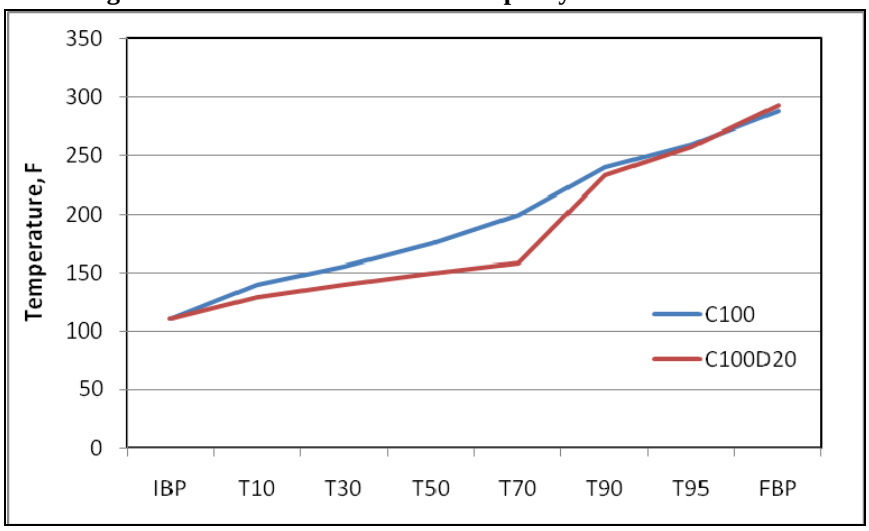


Figure 3.1: Distillation Curve Discrepancy for Fuel C100E20.



3.3 Role of Indolene Fuels in the Analysis

The AVFL-13 engine testing periodically included a series of tests on indolene fuel. During the first phase of the testing, the indolene fuel was run to "…determine the baseline engine operating conditions (valve timings, fuel-rail pressure, single injection timing, split injection timings and

quantities, and overall fueling quantities)…" These baseline engine-operating conditions were the starting conditions for testing each of the test fuels. Only the valve timings were changed to achieve the required operating conditions (primarily combustion phasing and engine load). Four repeatability tests with indolene were interspersed in the test fuel matrix to assure the engine operation remained consistent throughout testing." (Ref. 1, p. 3).

The testing was planned to measure engine performance under specified conditions:

- At fixed combustion phasing target defined by MFB50 of +5 degrees ATDC
- At power output targets for IMEP equal to 1.5 bar and 3.0 bar in Modes 1 (1000 rpm) and 2 (2000 rpm), respectively
- At a target of 5.5 bar for the maximum rate of pressure rise (MRPR) in Mode 3 (3000 rpm). This target results in IMEP values approximating 2.7 bar.

In addition, intake air temperature was subject to variation from the nominal values for each mode, consisting of ambient temperature (~25° C) in Modes 2 and 3 and of 100° C in Mode 1.

As a practical reality of engine testing, each test data point will deviate to some extent from one or more of these target conditions, and the measured engine performance will be influenced by the nature and extent of the deviations. Therefore, the comparison of engine performance between data points taken using different fuels or in different engine control modes will be influenced by this experimental imprecision, in addition to the effects of fuels or control modes that are the purpose of the comparisons. Some form of correction or control for experimental imprecision is needed to obtain the best possible assessment of fuel and control mode effects from the data.

For the Battelle analysis of the Phase 1 engine tests, the indolene repeatability tests were used to estimate a corrective model of engine performance versus observed deviations from the experimental targets, which was then used to normalize each engine test on the experimental fuels based on its own observed deviations from the targets. The corrective model involves terms for: the deviation from the MFB50 target; the deviation from the engine output target (either IMEP or MRPR depending on mode); an interaction between the MFB50 and engine output deviations; and the deviation of intake air temperatures from nominal values of 25° C or 100° C depending on mode. The coefficients for these terms measure the percentage change in engine performance per unit of deviation from the targets. To normalize the experimental data taken using non-indolene fuels, the corrective model is evaluated for the target deviations observed in each individual test; the percentage change in engine performance that would be expected for indolene is removed from the observed engine performance on the experimental fuel. This normalization method makes the fundamental assumptions that deviations from experimental targets will have the same percentage effects on engine performance for the experimental fuels as that which was observed for indolene and that sufficient indolene runs were performed to accurately track experimental variations.

The present analysis takes a different approach to compensating for the effect of experimental imprecision. The statistical corrections have been incorporated directly in the engine response models, where the coefficients of the statistical controls can be estimated simultaneously with the coefficients for the terms involving fuel and control mode effects. The fuels analysis indicates that indolene is outside the range of fuels that can be blended using the experimental blend stocks, particularly in having higher RVP and different chemistry at the front of the distillation

curve. Given these differences, the investigators preferred to estimate statistical controls within the framework of the engine response models using only the experimental fuels. Three terms are included to control for the deviations from the MFB50 target, the engine output target (either IMEP or MRPR depending on mode), and the nominal values for intake air temperature. No control was included for the interaction of MFB50 and engine output deviations. Section 6.2.1 describes the mathematical formulation of these control variables.

Predictions made for individual tests for comparison to the engine responses observed in the experimental data employ the values of the target deviations specific to the tests. When the models are used for general studies of engine behavior with respect to fuels and the control modes, these deviations are set to zero to remove the effects of experimental imprecision. By this method, general predictions of engine behavior made with the engine response models pertain to conditions that conform precisely to the experimental targets.

The continuation of engine testing in Phase 2 raises the question of the stability of engine performance over time. More than one year elapsed between the end of Phase 1 and the beginning of Phase 2, and the engine was then subsequently rebuilt between the Phase 2a and 2b testing. Accumulated wear on the engine from continued use coupled with maintenance and minor changes have the potential to introduce differences in the engine response to fuel and control mode variables between Phases 1 and 2, and the engine rebuild during Phase 2 is almost certain to do so.

The indolene testing in Mode 3 (3000 rpm) was used here in an exploratory analysis to determine the extent of these differences. The results indicated clearly that after the rebuild between Phases 2a and 2b, the engine ran more smoothly, was less smoky and emitted lower HC after the rebuild, while having later EVCA (Exhaust Valve Closing Angle) values, higher NO_x and longer measured combustion durations. Further, it was possible to detect differences between Phases 1 and 2. Using a weighted-average of Phases 1, 2a, and 2b as the baseline, the engine was appreciably smokier and emitted higher HC during Phase 1. Smoke and HC emissions were decreased modestly below the baseline in Phase 2a, and decreased much below the baseline in Phase 2b following the engine rebuild. Other engine response variables were found to differ among the phases. These differences by phase are seen clearly in results presented in Section 8.3.

The investigators chose to incorporate the statistical controls for phase differences into the engine response model in a manner similar to the controls for environmental imprecision. The method is to include a dummy variable representing the aggregate difference of Phase 2 from Phase 1 and a second dummy variable representing the difference between Phase 2a and 2b induced by the engine rebuild. Predictions made for individual tests for comparison to observed values employ the dummy variable values specific to the tests. Use of the models for general studies of engine behavior with respect to fuels and the control modes evaluate the dummy variable terms using the baseline defined by the number of tests conducted in Phases 1, 2a, and 2b.

4. FUEL MODEL

A statistical model of fuels was created in this work to support the analysis of fuel characteristics. PCA was used to identify generalized and independent vector features of the base experimental fuels using fuel chemistry and boiling points as independent variables. This vector representation of fuels is used in the fuel model to predict other physical properties of the fuels

and in the analysis of engine performance to represent the effect of fuels on engine performance. The vector representation was then extended to oxygenated fuels by adding ethanol content to the set of variables. Further analysis was conducted to determine how the addition of ethanol changes the properties of base fuels and how fuel ethanol content changes HCCI engine performance compared to the performance on hydrocarbon fuels. A total of 14 physical and chemical properties are available through the fuel model, either because they are part of the vector definition of a fuel or can be predicted from it.

4.1 Characterization of Experimental Fuels

In Phase 1, the ten experimental fuels (as described in Section 3.1) were chosen from the larger set of candidate fuels that could be blended from the four blending streams in a manner that was designed to minimize the degree of similarity and maximize the differences in terms of physical properties and chemistry among the fuels. The crux of the selection problem was stated as:

> "Pairwise plots of the 14 fuel properties revealed that several sets of fuel properties are so highly correlated that subsequent analyses of the relationships between engine performance and fuel properties could not distinguish between the effects of properties within each set. Thus, unless additional candidate fuels were prepared in a way that would reduce the correlation of these properties, it is necessary to choose only one of the properties to serve as a surrogate for all properties in the set." (Ref. 1, pg. 15)

This problem arises because motor vehicle fuels are complex hydrocarbon blends in which the conventional variables describing physical properties are not independent of each other, but rather are integrally tied together through the chemistry of the fuels. While it is common practice to use individual properties such as octane or boiling points to describe fuels used in research programs, it is much more difficult to ascribe the effect of fuels on engine performance to any one of the variables individually, or to any small subset of the variables. In some cases, it may be impossible to distinguish the effect of one variable from another because of the correlations that exist among properties. Even when the distinction is possible to draw, the individual variables being used are actually surrogates for other, more complex effects involving a number of different variables. When the correlations among variables exist because an underlying physical reality connects them and not merely because of experimental design, the conventional modes of analysis often have difficulty in determining causality.

PCA was used in Phase 1 to guide the selection of the ten experimental fuels (Ref. 1, Section 2.3). In this, PCA was used to transform seven chosen fuel properties into seven new, multivariate properties called principal components. Because PCA defines the principal components to be independent of each other, the principal components could be used as independent variables in classical experiment design techniques. A statistical measure was created for the distance of each fuel from its neighbors and was used in a selection process that maximized the distances among the ten selected fuels.

4.2 Development of Vector Fuel Model

While used initially to guide the selection of experimental fuels, PCA has been used more widely in the present work to create generalized, multivariate measures of the characteristics of the fuels. Because the correlations among octane levels, density and boiling points result from the underlying connection between fuel chemistry and physical properties, the vector variables

(principal components) created by PCA are viewed as more appropriate and useful measures of fuel characteristics than the original property variables. PCA has been applied in a number of ORNL studies related to fuels, as given in References 5, 6, and 7.

In particular, PCA is a solution to the problem of causality faced when using correlated variables. This is because they carry the correlations present in the data through the analysis without requiring artificial choices among surrogate variables at its start. All of the original variables contribute to the analysis through their role in defining the vectors themselves, thereby retaining the highest fidelity to the original data. The vector variables are also statistically efficient because, in circumstances where correlations are present among the original variables, there are fewer degrees of freedom (free choices) actually present in the data than the number of variables would suggest. The PCA representation will require fewer vectors to carry the systematic information on fuel characteristics.

Mathematically, PCA decomposes the correlations present among N variables in a set of data into multi-dimensional vector variables. PCA works on the basis of variables standardized to a mean of zero and standard deviation of one. If there are N variables, then the standardized variance in the dataset is also N. Having formed the matrix of independent variables in standardized form, PCA, in the form used here, performs a *singular value decomposition* of the correlation matrix to produce a slate of N eigenvectors and N eigenvalues. The eigenvectors are generalized, multi-dimensional descriptions of properties or characteristics found empirically in the data. The eigenvalues give the (standardized) variance associated with each eigenvector. Each entry in the dataset (here, each fuel) can be "scored" in terms of how much (or little) of the vector characteristics it expresses. The PCA representation has certain mathematical advantages in that its vector variables are independent of each other and produce variable scores that can be used as a new, uncorrelated set of variables. Further, the PCA representation usually involves a smaller number of vector variables that more closely match the degrees of freedom (unconstrained choices) actually present in the data.

The final result is to create a new coordinate system based empirically on the dataset in question. The origin of the coordinate system is located at the center of the cloud of data points, and the system of N orthogonal axes have been rotated with respect to the original variables to parse the variance in the data into orthogonal components. If the cloud of data points can be visualized as a football shape, the first eigenvector is oriented in N-space along the direction in which the data show the greatest variance – the long axis of the football. The second eigenvector is oriented at right angles along the direction in which the data show the next greatest variance, and so forth. If the football were elliptical in cross-section, the second principal component would be oriented along the major axis of the ellipse, and the third would be oriented along the minor axis. The eigenvectors define axis directions in N-space, and the eigenvalues measure the normalized variances along the axes.

4.2.1 Selection of Variables in Fuel Models

PCA was used to examine a number of different representations of the fuels during the course of the analysis. These representations included models based solely on physical properties such as octanes, density and boiling points, models based solely on fuel chemistry, and composite models based on both chemistry and properties. The objective was to find a representation that captured the fuel characteristics as best as possible, while relying on fuel chemistry as much as possible. Also considered was how to represent the ethanol content present in oxygenated fuels,

whether as a part of the PCA analysis of fuels or as a separate variable added for oxygenated fuels.

After consideration of alternatives, the representation shown in Table 4.1 was chosen as best meeting the objectives in the fuel model, although other representations could have been used with equal success. Nine variables describing fuel chemistry and boiling points were chosen to represent base (hydrocarbon) fuels as the variables in a PCA analysis. The six fuel chemistry variables define broad classes of chemical compounds, but do not provide a detailed speciation of the fuels. It was found that the chemistry variables did not fully capture the variation in measurable fuel properties. In particular, chemistry did not adequately predict boiling points, due most likely to the availability of chemical species within each of the five classes that have different distillation characteristics. Therefore, the boiling points T10, T50, and T90 were added to the variable set to complete the characterization of base fuels.

Table 4.1: Fuel Chemistry and Property Variables in Fuel Model (Independent and Dependent Variables)

Independent Variables for Base Fuels (used in PCA analysis)		
nParaXC_4	n-Paraffin content (excluding C_4)	volume %
nParaC_4	n-C_4 content	volume %
iPara	iso-Paraffin content	volume %
cycloPara	cyclo-Paraffin content	volume %
Olefins	Olefins content	volume %
Aromatics	Aromatics Content	volume %
T10	Temperature at which 10% distilled	° F
T50	Temperature at which 50% distilled	° F
T90	Temperature at which 90% distilled	° F
Added Independent Variable for Oxygenated Fuels		
EtOH	Ethanol Content	volume %
Dependent Variables Predicted by Fuel Model		
RON	Research Octane Number	number
MON	Motor Octane Number	number
RMSens	Sensitivity = RON – MON	number
SpGrav	Specific Gravity	gm/cm^3
RVP	Reid Vapor Pressure	Psi
T10 [a/]	Temperature at which 10% distilled	° F
T50 [a/]	Temperature at which 50% distilled	° F
T90 [a/]	Temperature at which 90% distilled	° F

[a/] Boiling points are dependent variables for oxygenated fuels only, based on the predicted change in boiling points of the base fuel (independent variables) caused by the added ethanol content.

The representation of oxygenated fuels was resolved in favor of adding ethanol content as a separate variable to the vector representation of the base fuel, rather than including ethanol content in the PCA-based fuel model. PCA is best applied in circumstances where the variables are inherently inter-related. This condition is not satisfied in the case of oxygenated fuels because ethanol can be added to any base fuel. Further, the use of splash blending, with attendant changes in octane and fuel volatility, causes the oxygenated fuels to form a distinct class. Therefore, a two-stage fuel model in which ethanol content modifies the characteristics of a pre-specified base fuel was found to be statistically preferred and in keeping with the nature of splash blending.

The independent variables {nParaXC_4, nParaC_4, isoPara, cycloPara, Olefins, Aromatics, T10, T50, and T90} define a base fuel. The PCA analysis of these nine variables yields nine vectors as an alternative description of base fuels. A subset of the vectors were chosen to represent the systematic information regarding base fuels and then used as predictors for HCCI engine performance. As a second step, ethanol content is added to the selected vectors representing base fuels to create a fuel model capable of representing both hydrocarbon and oxygenated fuels. The effect of fuel oxygenation on engine performance was determined using a paired analysis in which performance on oxygenated fuels is compared to performance on the corresponding base fuels.

The fuel model is completed by its use in predicting the other measured properties of the fuels, including RON, MON and sensitivity, specific gravity, RVP, and the final boiling points of oxygenated fuels. This also is a two-stage process. In the first stage, statistical models are developed to predict RON, MON, sensitivity, specific gravity, and RVP as functions of the vector representation of the base petroleum fuels. In the second stage, additional statistical models are developed to predict how added ethanol addition changes the properties determined for the base fuel as a function of the ethanol amount blended. In an oxygenated fuel, the fuel chemistry variables are diluted from the base fuel values, and the octanes, density, volatility and boiling points are modified from the base fuel value based on the predicted effects of ethanol. A total of 14 different variables are available through the complete fuel model to describe the chemistry and physical properties of the experimental fuels.

4.2.2 Development of Vector Model for Base Fuels

An attractive feature of the vector variables created by PCA is *dimensionality reduction*. When the correlations among variables are large, the degrees of freedom actually present in the data is often much smaller than the number of variables in the dataset. In this case, the result of applying PCA to the dataset of nine variables for the base fuels yields nine vectors, of which only three to five need be carried to characterize the systematic information about the fuels. Table 4.2 shows the partitioning of the fuel variance by eigenvector; the fuel variance is a statistical measure of the extent to which the fuels differ from each other in terms of the chemistry and boiling point variables that make up each of the vectors. Another way of saying this is ‘what % of the fuel range tested can be explained by each vector?’ Vector 1 is associated with 47% of the variance present in the fuels dataset, followed in order by Vector 2 (36%) and Vector 3 (13%). Together, the first three vectors represent more than 95% of the variance in the data. They are followed by a smaller Vector 4, representing an additional 3% of the variance and bringing the cumulative variance up to nearly 99%. Vector 5 (1%) would be required to bring the cumulative variance above 99% and an additional four vectors would required to capture 100%.

Table 4.2: Partitioning of Fuels Variation by Principal Components.

Principal Component	Eigenvalue	Percent of Variance	Cumulative Percent
1	4.24	47%	47.1%
2	3.20	36%	82.7%
3	1.17	13%	95.7%
4	0.26	3%	98.6%
5	0.06	1%	99.3%
6	0.05	1%	99.8%
7	0.01	0%	99.9%
8	0.00	0%	100.0%
9	0.00	0%	100.0%
Total	9.00	100%	-

PCA produces as many eigenvectors as there are variables in the dataset as a mathematical necessity, so that the eigenvectors form a complete vector basis that can be used to exactly reproduce the original data. When PCA is used because the original variables are strongly correlated, it is usual for a subset of vectors to capture nearly all of the variation in the data. Unfortunately, there is not a universal method for making a judicious choice of vectors. In general, vectors with eigenvalues of at least 1 are retained, because each of the original variables contributed a normalized variance of 1. Vectors of similar size contribute information equal to a "full" variable, and a minimum of three vectors must be retained for this fuel set based on this criterion. One may also consider the number of vectors required to reach a given cumulative level of variance explanation; four vectors would capture almost 99% of the variation among fuels.

In this case, the best guidance comes from knowledge that finding three major and one minor eigenvectors is not an accident, but is rather a reflection of the degrees of freedom in the data:

- The three major degrees of freedom were introduced by the blending substitutions of A for D; B for D; and C for D in creating the base fuels
- An additional, minor degree of freedom was introduced by the re-sampling of blend stream C during Phase 2 of the study, introducing a blend stream C' with slightly different chemical and physical properties.

The first four eigenvectors are carried in the development of the engine performance and fuel models, capturing a cumulative 99% of the variation among fuels. In general, the predictive models employ one or more of Vectors 1, 2, and 3 to represent fuel effects, while an occasional contribution is made by Vector 4.

4.2.3 Definition of Fuel Vectors

The principal components defined in PCA are vectors in N-space that form the axes of a new coordinate system. Each vector consists of a linear combination of the original variables, where the variables are expressed in normalized form and internal coefficients give the relative weights of each of the variables in defining the vector. In applying PCA to the fuels data, the six chemistry variables were first transformed as described in Appendix A in order to implement the physical constraint that compositional variables can take on values only in the interval from zero to one. The transformation is one that maintains a nearly-linear relationship to the original chemistry variables over the range of the data; its use is indicated by the prefix trn as in trn(ParaXC4). The three boiling point variables were added to complete the list of variables. All of the 15 base fuels used in the analysis are contained in the dataset, including separate entries for the fuels re-blended or re-measured for Phase 2.

Table 4.3 presents the first four vectors of the fuel model as output by PCA. Vector 1 can be read in a literal sense as saying that the first principal component axis lies in a direction in space along which the variable trn(ParaXC4) increases by 0.269 standard deviations (s.d.), while trn(ParaC4) increases by 0.372 s.d, trn(iPara) changes by 0.02 s.d., and so forth. The mathematical equation representing any Vector i is:

$$V_i = c_{i,1} * \text{trn(ParaXC4)} + c_{i,2} * \text{trn(ParaC4)} + c_{i,3} * \text{trn(iPara)} + c_{i,4} * \text{trn(cycloPara)}$$

$$+ c_{i,5} * \text{trn(Olefins)} + c_{i,6} * \text{trn(Aromatics)} + c_{i,7} * \text{T10} + c_{i,8} * \text{T50} + c_{i,9} * \text{T90} \quad \text{(Eq 4-1)}$$

In this formalism, the coefficients $c_{i,1}$ through $c_{i,9}$ are the internal coefficients for the vector as given in the table, and the variables defining the vector are used in the calculation in their normalized form (standardized to mean 0 and standard deviation 1). When Eq. 4-1 is applied to the values for a specific fuel, the computed values V_i are the vector "scores" for the fuel, which measure the extent to which the fuel is similar (or dissimilar) to the characteristic represented by the vectors. The vector scores are the values used in this analysis as alternative descriptions of the fuels.

While the meaning of the vectors is complicated by the mathematics, what is clear, at least directionally, is that the vectors represent the following trends:

- Vector 1 – a trend toward fuels with more n-paraffins and aromatics and less olefins, leading to elevated distillation curves compared to the average fuel. In the opposite direction along this axis, one would find fuels with more olefins and less n-paraffins and aromatics and lower distillation curves.
- Vector 2 – a trend toward fuels with more n-paraffins (excluding C_4), cyclo-paraffins, olefins and aromatics, and with less C_4 and iso-paraffins, along with varied minor effects on the distillation curve compared to the average fuel.
- Vector 3 – a trend toward fuels with more n-paraffins and cyclo-paraffins, and slightly more iso-paraffins and olefins, with less aromatics; the distillation curve is made steeper primarily through depression of T10.
- Vector 4 – a trend toward fuels with less n-paraffins (excluding C_4) and iso-paraffins, and more of the remaining constituents, resulting in a much steeper distillation curve that is depressed at T10 and elevated at T50 and T90.

Table 4.3. Definition of Vectors Forming Fuel Model.

	Vector 1	Vector 2	Vector 3	Vector 4
Percent of Fuel Variance	47%	36%	13%	3%
Cumulative Percent	47%	83%	96%	99%
Internal Coefficients				
trn(ParaXC4)	0.269	0.427	0.265	-0.290
trn(ParaC4)	0.372	-0.246	0.393	0.030
trn(iPara)	0.020	-0.552	0.098	-0.143
trn(cycloPara)	0.034	0.387	0.654	0.115
trn(Olefins)	-0.422	0.253	0.092	0.175
trn(Aromatics)	0.208	0.376	-0.544	0.189
T10	0.445	0.059	-0.156	-0.631
T50	0.448	0.178	-0.083	0.360
T90	0.410	-0.251	0.042	0.536

Because the meaning of the vectors is difficult to intuit from the tabulation of their internal coefficients, an interactive tool called the Fuel Panel has been created to display the effects of the vectors in a more intuitive manner. Section 9 describes its use, along with the other simulators created in the work.

As seen above, the vectors represent changes in the chemistry of fuels, with corresponding effects on the distillation curve. While generalized by PCA to have desirable mathematical properties, the vectors are actually quite similar to the blend stock substitutions that were used to create the experimental fuels. Any of the substitutions – for example, a change from A100 to A90D10 – induces a change in the chemical composition and in other properties of the fuel, including boiling points. The changes are proportional to the volumetric extent of the substitution, but occur in fixed ratios among the variables. The same qualitative properties hold for the vectors of the fuel model, but the use of variables in normalized form (rather than physical units) tends to obscure the similarity.

Figure 4.1 shows the similarity between vectors and blend stream substitutions by plotting the vector scores for the experimental fuels in two-way scatter plots. Vector scores are measures of the degree to which the fuels express the characteristics represented by the vectors; they are easily calculated from the fuel properties and the internal coefficients of the vectors as demonstrated in Appendix A. A score is positive when the fuel deviates from the average fuel in one direction along the axis and is negative when the deviation from average is in the opposite direction. In the first case (positive scores), one might say that a fuel is much like the characteristic represented by the vector, while in the other case (negative scores) one might say that the fuel is much unlike the characteristic. While "score" may seem an abstract term, it is simply a measure of where a fuel falls along a principal component axis. It is no different than saying that a particular fuel has a score (value) of 89 as measured along the octane axis.

From the figure, we see that Vector 1 represents a characteristic that is explored by the A vs. D and C vs. D blending substitutions (in addition to the single fuel A50C20D30 involving all three blend streams). As the quantity of D is increased from zero, the A-series fuels move from right to left along the Vector 1 axis at a nearly-constant value for Vector 2. As the quantity of D is increased in the C-series, the fuels move from left to right along the Vector 1 axis at a nearly-constant Vector 2 score. Thus, Vector 1 represents the tradeoff of chemistry and its related effect on boiling points that is explored by the A-series and C-series substitutions. Although the two series do not explore the characteristic represented by Vector 2, they are distinguished from each other by their differing locations on the Vector 3 axis. As seen in the lower portion of the figure, the A-series fuels generally score low on the Vector 3 axis, while C-series fuels score higher. The fuels designed from B vs. D substitutions mix together the Vector 1 and 2 characteristics and follow an upward sloping line from left to right as the amount of D in the blend increases from zero. The B-series fuels vary along similar diagonals in the other portions of the figure.

The fuel model vectors are seen to be generalized representations of the changes in fuel chemistry and properties that were created in the experimental fuels through the systematic substitutions of blend streams A, B, C and D. The vectors are not precisely aligned with the blend stream substitutions, but have been rotated in N-space by the PCA procedure to obtain the desirable mathematical properties of being independent variables over the dataset. Each vector

represents a particular change in fuel chemistry and boiling points that was explored in the design of the fuels.

The eigenvalues state the proportions of the total variation in fuels associated with the vectors and, therefore, give a measure of how much of the experiment is devoted to each of the fuel changes. The fact that the first three vectors have a dominant role in explaining the fuel variation is a reflection of the use of three substitution strategies (A vs. D, B vs. D, and C vs. D) in blending the fuels. The presence of a fourth, smaller vector is a result of the re-sampling of C in Phase 2 of the program. Vectors 1 through 4 do not represent the three blending substitutions and the C re-sampling individually. Rather, they are present in these numbers and relative sizes because of the degrees of freedom known to be present in the experimental fuels.

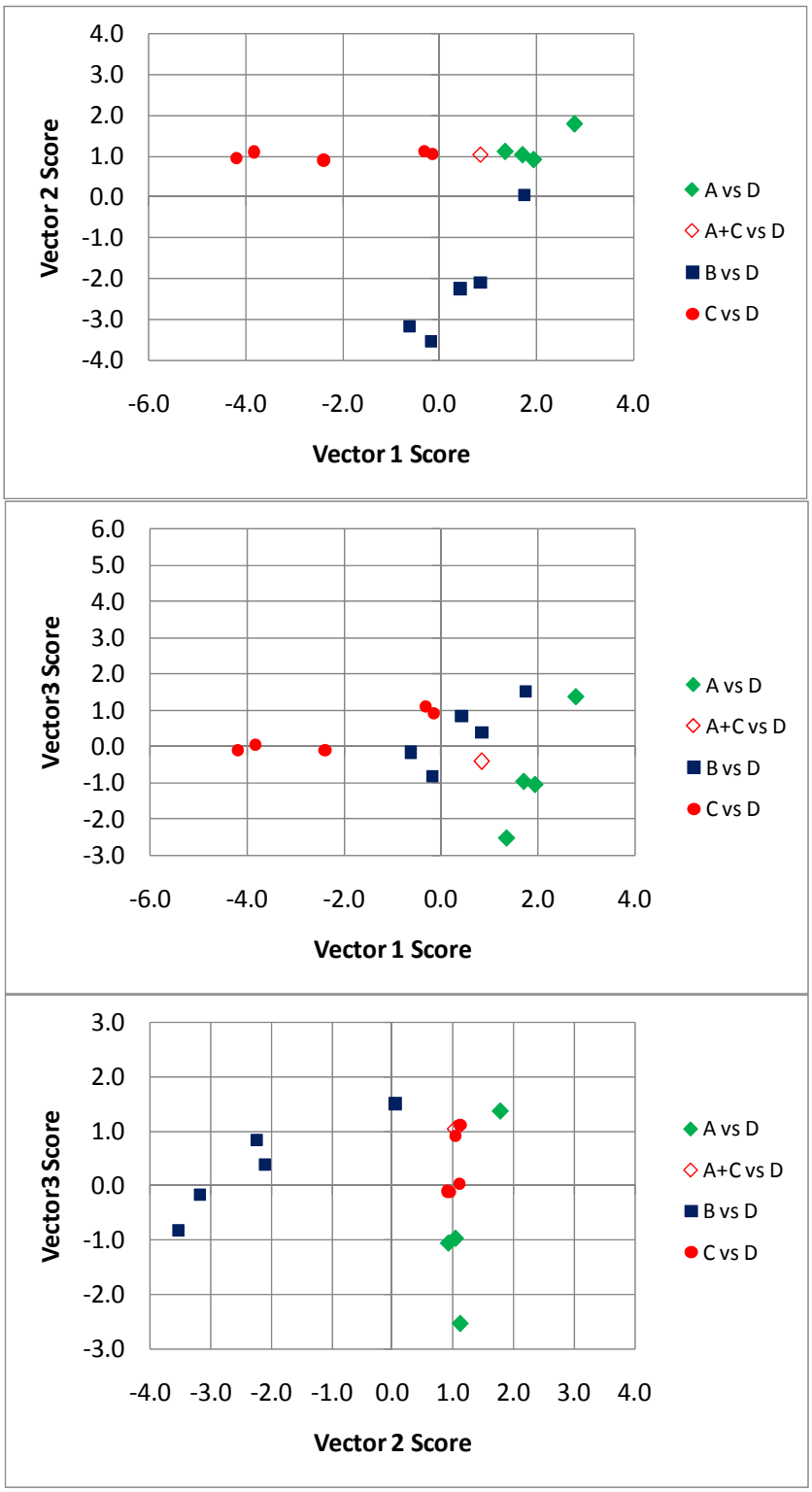


Figure 4.1: Distribution of Experimental Fuels in Vector Space.

4.3 Statistical Model of Fuel Properties

One aspect of the fuel model is the ability to predict additional physical properties of fuels, given the set of chemistry and boiling point variables used to define a base fuel and the presence, if any, of ethanol in the blend. The following sections describe the predictive models of fuel properties.

4.3.1 Base Hydrocarbon Fuel Properties

A statistical model of the properties of the base hydrocarbon fuels was developed using conventional regression analysis in which Vectors 1 through 4 of the fuel model were the independent variables and the measured properties of the 15 experimental fuels – including RON, MON, RMSens, SpGrav and RVP – were the dependent variables. The boiling points T10, T50, and T90 were included as dependent variables for completeness in the base fuel property models, although they are not true dependent variables because they are used in the definition of the vectors. The forms of the predictive equations and the coefficient values are contained in the fuels simulator, which can be exercised to make predictions for user-specified fuels and also serve to document the predictive equations.

As shown in Figure 4.2, the models do a very good job of replicating base fuel properties, with the R^2 statistic ranging from 0.95 to 0.99, except for RVP. Vectors 1, 2 and 3 are statistically significant predictors for RON, MON, RMSen and SpGrav, while Vector 4 is not statistically significant in any instance. The tendency for residuals in RON and MON to be consistent in direction (see open circles in figure) indicates the presence of a common, but unmeasured, influence. Because of the common influence, the model for sensitivity (RMSens) gives better predictions than the models for either RON or MON. The models for T10, T50, and T90 necessarily involve all four vectors because the boiling points are used to define the vectors and all of the vectors contain information about the boiling point variation of the fuels.

Although RVP was controlled to a nominal 7 psi target, the measured RVPs varied from 6.6 to 7.5 psi. An effort was made to develop a predictive model for the RVP variation within this restricted range, but without success. Vectors 1-4 are found to have no predictive power for RVP. Because RVP was controlled to its target by the addition of n-butane to a fuel blend, a second effort was made to identify the amount of n-butane added to the blend, as distinct from the total n-butane present in the final fuel, but this also had no success as a predictor for RVP. Conceptually, the measured RVP value should be determined by the chemical composition of the fuel, although no variable in the dataset contains the information needed to make such predictions. This null result is not fully understood, but it may suggest that random variation in laboratory measurement of RVP swamp any systematic effect that fuel specifications have on the RVP variation seen in the experimental fuels.

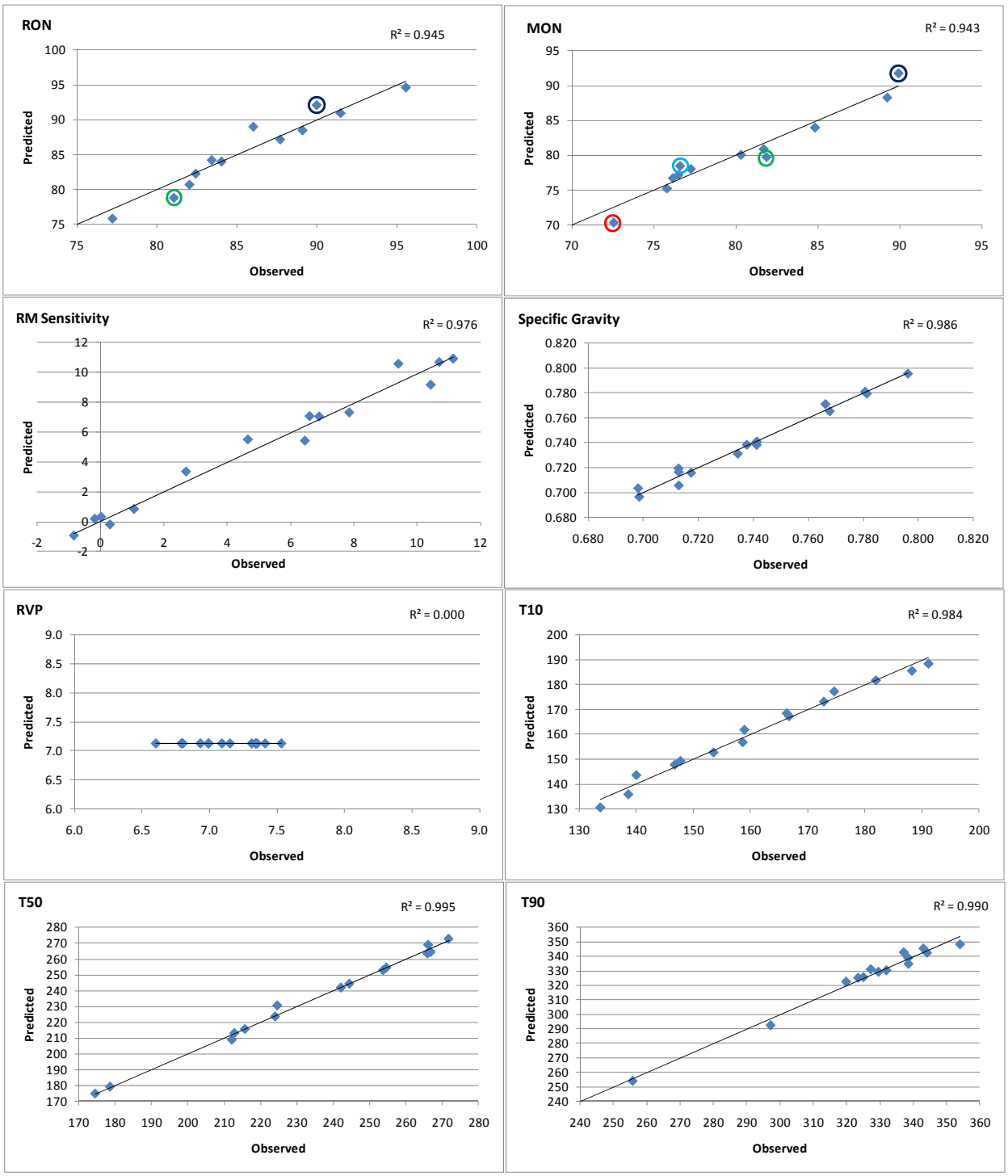


Figure 4.2: Predictive Power of Base Fuel Property Model.

4.3.2 Statistical Model of Ethanol Effects on Fuel Properties

Phase 2 of the program splash blended ethanol into selected base fuels, in target amounts varying from E10 to E30. The addition of ethanol dilutes the base fuel volumetrically and causes a number of fuel properties to change as a function of the quantity of ethanol added. To represent

these effects in the fuel model, a paired study was conducted in which the properties of each oxygenated fuel were compared to the properties of the base fuel from which it was blended. Figure 4.3 gives an example of this pairing.

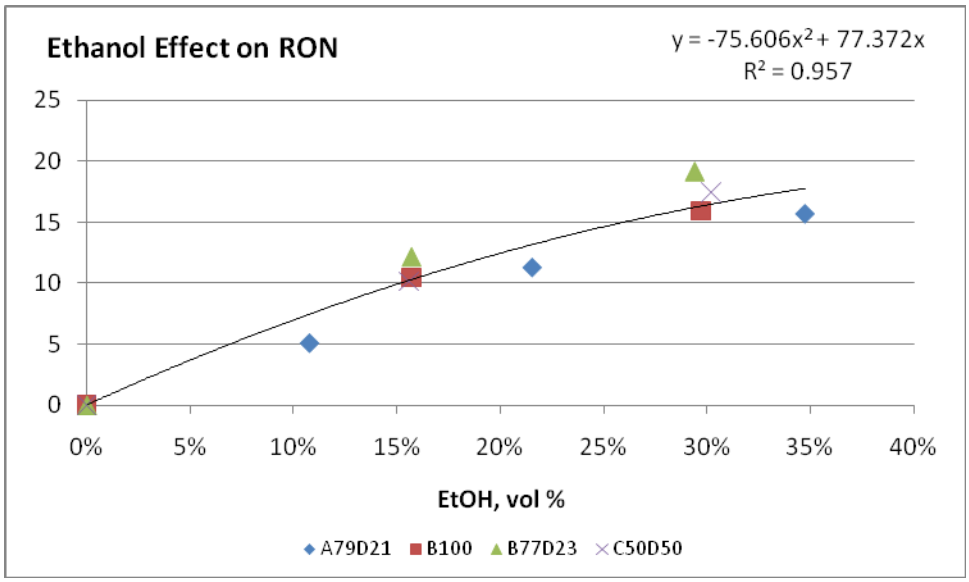


Figure 4.3: Example of Paired Analysis for Ethanol Effects on Fuel Properties.

In this analysis, ethanol content (volume %, as measured) was the primary independent variable, while the variables RON, MON, RMSen, SpGrav, RVP, T10, T50, and T90 were the dependent variables. In this instance, the boiling points are true dependent variables because the presence of EtOH changes the distillation curve. The model forms allowed for a quadratic term in ethanol content, in addition to the linear effect, and interactive terms with base fuel properties. The properties of oxygenated fuels are found to vary with ethanol content, showing evidence of non-linear effects in some instances, but only relatively small interactions with properties of the base fuel (MON, density and the boiling points). As for the model of base fuel properties, the forms of the predictive equations and the coefficient values are contained in the fuels simulator, which can be exercised to make predictions for user-specified fuels (as described in Section 9) and also serve to document the predictive equations.

Figure 4.4 shows the primary effects of ethanol blending as predicted by the fuel model. Octane, both RON and MON, are shown to increase with the addition of EtOH, but at rates that show diminishing returns at higher oxygenate levels. RON increases more rapidly than MON, so that sensitivity increases significantly. Although the fuel model is not able to predict the actual RVP of individual base fuels, it does show the expected trend of RVP with the addition of ethanol.

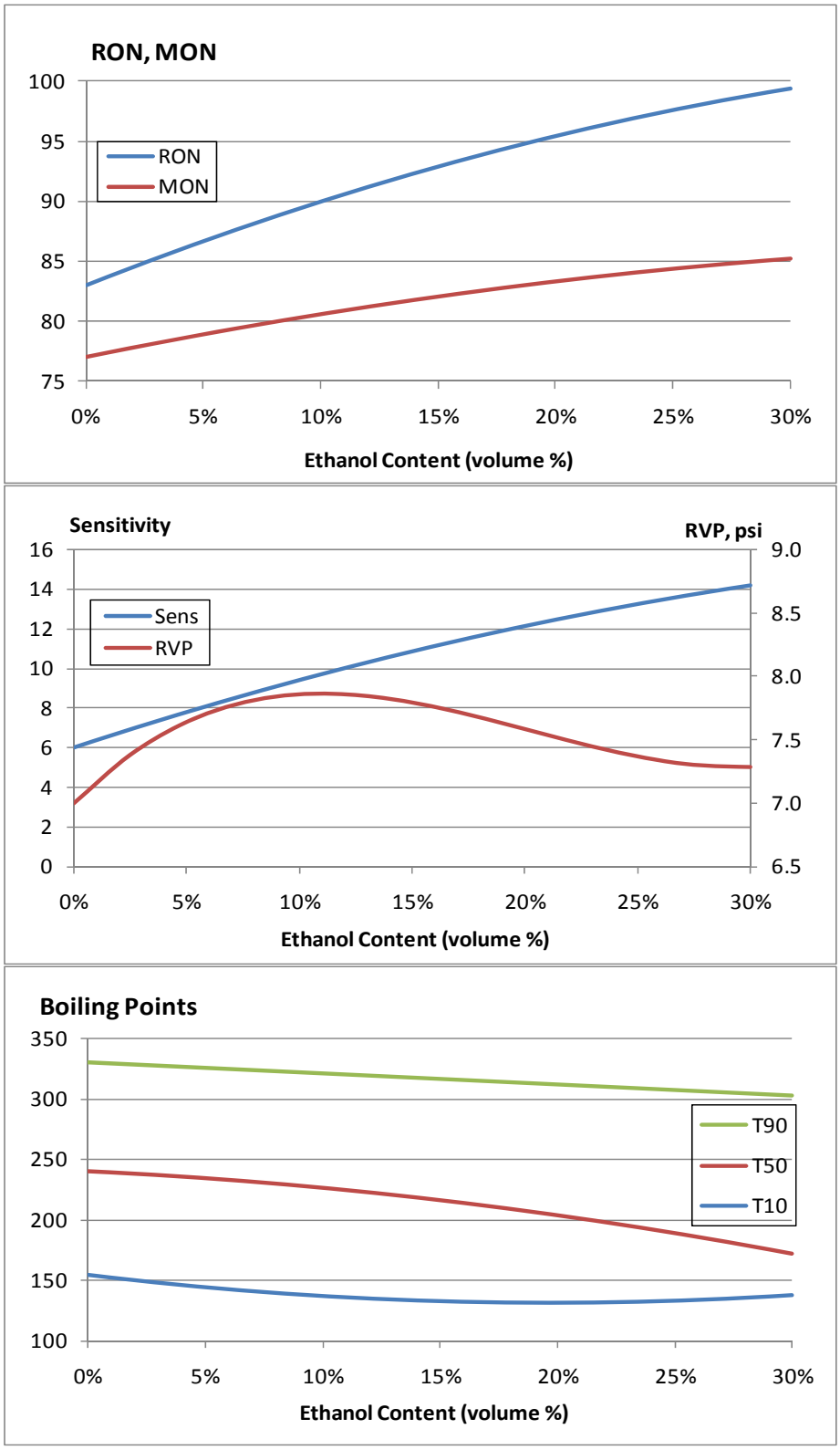


Figure 4.4: Primary Effects of Ethanol Blending for the Average Fuel.

The RVP effect peaks at about 10% ethanol content, having increased RVP nearly 1 psi above the level of the base fuels, before turning around and beginning to decline back toward 7 psi at higher concentrations. This statistical result is consistent with the known chemical behavior of the ethanol molecule in hydrocarbon fuels. The addition of ethanol reduces all three boiling points below the levels of the base fuel. The effect on T10 levels off at higher concentrations, and the effect is greater on T50 than on T90, as expected.

Figure 4.5 illustrates the predictive ability of the complete fuel model. Base hydrocarbon fuels are plotted using blue diamonds, while the oxygenated fuels are plotted using red diamonds. With the exception of RVP, the model is generally successful in predicting the properties of the experimental fuels. The model has a good ability to predict octane, sensitivity and specific gravity. It has a generally good ability to predict boiling point effects, although the T50 and T90 predictions have substantial errors for some fuels. The model's ability to predict RVP is poor – it is unable to predict the variation in RVP for base hydrocarbon fuels, but it does capture the nonlinear increase in RVP with added ethanol content.

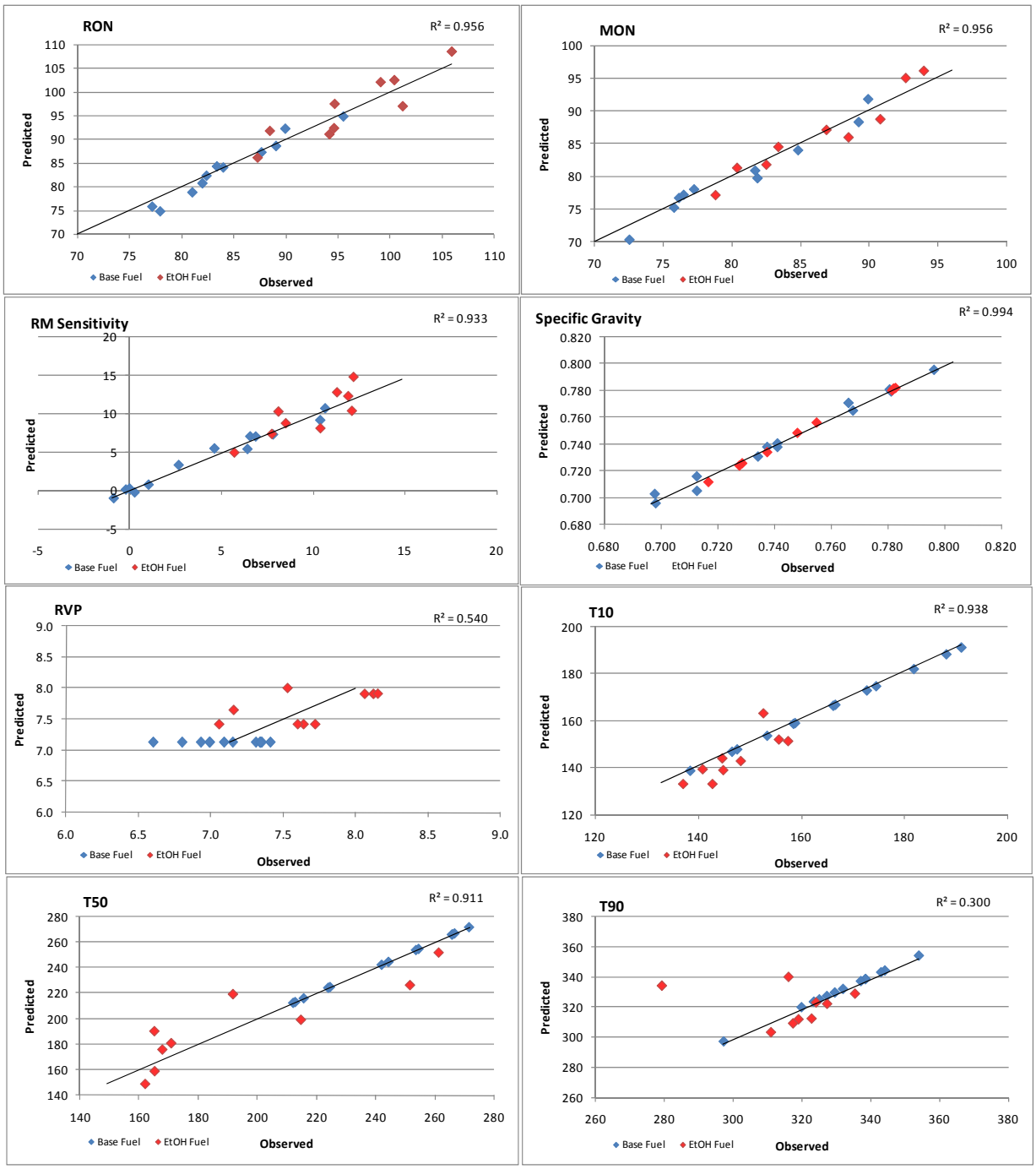


Figure 4.5: Predictive Ability of Final Fuel Model for Base and Oxygenated Fuels.

4.4 Generation of Parametric Fuels

The studies of engine performance presented in later sections are based on the analysis of engine response to a group of fuels generated parametrically to span a wide range. The term *parametric* is used here to mean one or more factors that determine the performance of a system and can be varied in an experiment. The factors of interest are the characteristics of fuels as represented by the vectors in the fuel model. As shown at the beginning of this section, the experimental fuels

outline the boundaries of a region in fuel space (see Figure 4.1). The purpose of generating parametric fuels is to fill the interior of the outlined space with a range of intermediate fuels that could have been produced by the same blending strategy. The fuels are then used to exercise the predictive models for engine response to identify the most important trends in engine performance with respect to fuel characteristics.

In concept, there are a number of different ways that a parametric fuel set could be generated. For example, one could construct a fuel model and parametric fuels directly from the measured properties of the blend Streams A, B, C, C', and D. Using an assumption of linear volumetric blending, for example, it would be possible to generate parametric fuels by selecting values for coefficients a_i, b_i, c_i, and c_i' in the equation below:

$$F_i = a_i A + b_i B + c_i C + c_i' C' + (1 - a_i - b_i - c_i - c_i') D \qquad \text{(Eq. 4-1)}$$

The coefficient values could be selected in either of two basic ways:

- Systematically, by setting discrete levels for each coefficient and then enumerating all possible combinations of levels of the four coefficients. The number of parametric fuels generated by this method is calculated as the product of the number of discrete levels assigned to each of the coefficients.

- Randomly, by making independent and random choices for the values of each of the four coefficients within a range of values determined to be suitable for the sampling. This method depends upon random chance to create combinations of the blend streams across the full extent of the fuel space and is known as random balance. Any number of parametric fuels can be generated by this method by repeated sampling.

Both methods will consider not only the two-way substitutions used to create the experimental fuels (i.e., A vs. D, B vs. D, C vs. D.), but also blends involving three, four or five of the blend streams. The first method does so systematically using combinations of discrete levels for each stream, but will leave gaps unexplored between the selected fuels. The second method does so according to the principle of random balance, but in small samples one cannot be confident of having evenly and equally explored all parts of the fuel space and all combinations of the blend streams. The analysis in the current work follows an approach for the generation of parametric fuels similar to that described above, but based on sampling from Vectors 1, 2, 3, and 4 of the fuel model. Both systematic and random sampling techniques have been used, but systematic sampling has generally been preferred because it assures even and equal coverage of the possible combinations in studies of modest size.

Most of the analysis on engine performance that is presented in this report is based on a parametric fuels dataset having 249 members. This set was generated using the fuels simulator described in Section 9 by sampling the values of Vectors 1 through 4 over a range of ± 1 standard deviation from the average fuel, and using from three to five discrete levels per vector. While PCA defines its vectors to be independent of each other, the actual fuel data used to define the vectors will not contain all combinations of the vector values that are possible. Thus, sampling the vectors can lead to combinations of vector values not seen in the actual data and to fuels with chemistries or properties outside the range seen in the data. Such fuels are possible

realizations of the blending strategy that produces the experimental fuels, but they lie outside the boundaries explored by the experimental fuels. Whether to include or exclude such fuels is a decision that must be made by the analyst in each case. For the parametric fuels set, fuels outside the range of experimental properties and chemistries were excluded, reducing the number of parametrically generated fuels to the 249 that survive. The fuel sets generated for parametric studies are discussed more fully in section 8.1.

5. EXPERIMENTAL ENGINE DATA

The engine used for the AVFL-13 and 13b test programs was a single cylinder research engine, loosely modeled after a European Ford passenger car engine, with a custom cylinder head equipped with two intake valves, one exhaust valve, and Sturmann hydraulic variable valve actuation hardware. Only one intake valve, equipped with a high swirl intake port, was used. The engine is also equipped with both port and direct fuel injection. Nominally, the engine was controlled slightly lean of stoichiometric air fuel ratio and combustion phasing was set to 5 degrees after top dead center for all test points. As summarized in Table 5.1, three operating points were run, 1000 rpm (1.5 bar IMEP), 2000 rpm (3.0 bar IMEP), and 3000 rpm (5.5 bar/deg peak pressure rise). For Mode 3, the 5.5 bar/deg rate of cylinder pressure rise corresponds to about 2.7 bar IMEP. The first point was chosen to represent idle, taking into account that the engine would not run well at either lower loads or lower speeds. The second point represents in-city cruise conditions. The third condition represents higher load, higher speed engine operation, but was limited by a maximum allowed rate of cylinder pressure rise for the engine.

Three engine control modes were run, consisting of port fuel injection with negative valve overlap (known as recompression-early injection or RCEI), port fuel injection with a second exhaust valve event during the intake stroke (known as re-breathing-early injection or RBEI), and direct injection with a second injection during the negative valve overlap period (known as recompression-split injection or RCSI). These three control modes were chosen to represent different options for achieving HCCI in order to determine if they had different operating characteristics or response to fuel changes. In the engine test program, the RCSI mode was not tested in Phases 2a or 2b.

Experimental variation arose from several defined sources. First, operating conditions such as temperatures, IMEP, or MFB50 were not set exactly for all experiments. And secondly, the experiments were conducted in three separate series of tests with engine and other changes between them.

Table 5.1: Definition of Engine Operating Modes.

Test Condition		Mode 1	Mode 2	Mode 3
Engine RPM		1,000	2,000	3,000
Combustion Phasing (MFB50, degrees ATDC)		+5 deg	+5 deg	+5 deg
Engine Output	IMEP	1.5 bar	3.0 bar	
	dPdCA			5.5 bar/deg
Intake Air Temperature (°C)		+100°	+25°	+25°

6. ENGINE SIMULATOR MODELS

Statistical models for the response of engine performance to fuels and control modes were developed independently for each operating mode, and then were used to build engine simulators in the form of Excel spreadsheets. Fuels are represented in the response models using the vector representation of fuels developed with PCA. Differences in engine response by control mode are represented using the RCEI mode as a baseline and introducing dummy variables for the RBEI and RCSI modes. The predictive models include statistical controls for imprecision in meeting the experimental targets for combustion phasing, engine output, and intake air temperature, and additional statistical controls for systematic differences in engine response among Phases 1, 2a, and 2b of the testing.

6.1 Development of Engine Models

The performance of the test engine in any given test can be thought of as a response to a number of variables:

- Engine speed (rpm), which also includes a corresponding load related target, as set by the dynamometer
- The engine control mode chosen, whether RCEI, RBEI, or RCSI
- The characteristics of the fuel chosen
- The combustion phasing achieved, measured by MFB50
- Engine output (as measured as IMEP or another variable), based on the fueling rate and other variables
- External variables that influence engine performance, such as the intake air temperature
- Internal variables related to changes in engine condition, such as those leading to differences in engine performance across the phases of the AVFL-13 testing.

In a testing program that measured engine performance with respect to all of the variables, including sweeps in combustion phasing and engine output, it should be possible to develop a

single predictive model to represent a wide range of engine operating conditions. The AVFL-13 testing takes a different approach, however, in that it measures engine performance only in three distinct operating modes, representing three well-separated regions in the full space of engine operations. A pilot study attempted to develop a single, integrated model spanning all three operating modes. While the model resolved the gross differences among the modes, it proved to have limited ability to resolve differences within mode related to the effects of fuels and engine control modes. Thus, the work performed here takes the approach of estimating independent models of fuel and control mode effects for each operating mode. The model structure is identical across the modes, but the coefficient estimates and the terms retained in the final models will differ across the modes.

Both Phases 1 and 2 of the engine testing considered the effect of varying characteristics of base (non-oxygenated) fuels on engine performance, while the effect of fuel oxygen was considered in portions of the Phase 2a and 2b testing. The AVFL Committee expressed an interest in having a single fuel and engine model that applied to both base and oxygenated fuels. Regression analysis is capable of resolving multiple causal effects that may be present in data, but a necessary condition is that the form of the regression models accurately fit the effects present in the data. In the context of this study, one must accurately specify how base fuel characteristics and fuel oxygen content affect engine performance in order to achieve the correct attribution of fuel effects. Any misspecification can lead to incorrect attribution of the effects between base fuels and fuel oxygen content. Because the oxygenated fuels were tested only in Phases 2a and 2b, and because fuel characteristics were not controlled during the splash blending, the Committee accepted the recommendation to separate the analysis of engine response between base and oxygenated fuels.

The analysis of engine response to base fuel characteristics was conducted using conventional multivariate linear regression, as described in the following section. The analysis of engine response to fuel oxygenation was conducted in a separate analysis that paired engine performance on oxygenated fuels to the performance on the corresponding base fuels, as described in Section 6.3.

6.2 Engine Models for Base (Non-Oxygenated) Fuels

6.2.1 Mathematical Formulation

The predictive models for engine response take the form of regression equations that relate engine performance, as measured by one of 15 dependent variables, to the conditions of the engine test, including engine control mode and base fuel characteristics. The engine performance metrics P_i are listed in Table 6.1, which gives the variable names, the units of measure, and identifies which variables were subject to a logarithmic transform. Dependent variables were entered in the statistical analysis as logarithms when their values spanned two or more orders of magnitude, when measured values could be close to zero but could not be zero or negative, or when the reported values are calculated as ratios of other variables. It is common practice to transform dependent variables in these circumstances to better represent what is often a log-normal distribution of values. Dependent variables were used in the physical units as reported when values were sufficiently far from zero to mitigate the need to avoid zero or negative values, when the range spanned was less than two orders of magnitude, or when they already were measured in log units (noise measured in dB). There are no universal rules for the decision to

transform the dependent variable, and different choices could be made in some instances; it is unlikely that different choices would lead to substantially different predictions.

Table 6.1. Summary of Engine Performance Metrics

Performance Metric	Variable	Units	Transform
Air-fuel ratio	lambda	number	ln
Exhaust valve closing angle	EVCA	degrees ATDC	none
Indicated Specific Fuel Consumption	ISFC	gm/kwh	ln
Indicated Thermal Efficiency	ITE	number	ln
Indicated Specific Hydrocarbons	ISHC	gm/kwh	ln
Indicated Specific Carbon Monoxide	ISCO	gm/kwh	ln
Indicated Specific Nitrogen Oxides	ISNOx	gm/kwh	ln
Filter Smoke Number	Smoke	number	ln
Noise	Noise	dB	none
Indicated Mean Effective Pressure	IMEP	bar	none
Maximum Rate Cylinder Pressure Rise	dPdCA	bar/deg	ln
Combustion Duration	CombDur	millisecond	ln
Peak Cylinder Pressure	PCP	bar	ln
Coefficient of Variation, IMEP	COV	percent	ln
Combustion Efficiency	CombEff	number	ln
Exhaust Temperature	ExhT	° C	none

Using the notation $P_{i,j}$ to denote the i-th performance metric P_i (transformed or not) as measured on the j-th engine test, the mathematical form of the engine response model for any operating mode can be written as follows:

$$P_{i,j} = a_i + b_i * (MFB50_j - MFB50_{target}) + c_i * (EngOutput_j - EngOutput_{target}) + d_i * (IntakeT_j - IntakeT_{target})$$

$$+ e_i * dPhase2 + f_i * dPhase2b$$

$$+ g_i * dRBEI + h_i * dRCSI$$

$$+ i_i * V_1 + j_i * V_2 + k_i * V_3 + l_i * V_4$$

$$+ dRBEI * (m_i * V_1 + n_i * V_2 + o_i * V_3 + p_i * V_4)$$

$$+ dRCSI * (q_i * V_1 + r_i * V_2 + s_i * V_3 + t_i * V_4) \qquad \text{(Eq. 6-1)}$$

While Eq. 6-1 is complex, containing 19 individual terms, not all terms are retained in the final models for the operating modes and the terms can be organized into easily understood groups:

- Coefficients b_i, c_i, and d_i are the statistical controls for the deviation of experimental conditions for test j from the nominal targets by operating mode given in Table 6.1.
- Coefficients e_i and f_i are the statistical controls for the systematic variation in engine response among the phases. The control variable dPhase2 takes the value 0 for tests conducted in Phase 1 and the value 1 for tests conducted in Phase 2. The control variable dPhase2b takes the value 0 for tests conducted during Phase 2a (before the engine rebuild) and the value 1 for tests conducted during Phase 2b (after the rebuild).
- Coefficients g_i and h_i measure the difference in the engine response in the RBEI and RCSI control modes, as additive corrections to the baseline engine response in the RCEI mode.
- Coefficients $\{i_i, j_i, k_i, l_i\}$ measure the engine response to the fuel characteristics represented by Vectors 1 through 4 of the fuel model. The coefficients $\{i_i, j_i, k_i, l_i\}$ represent the baseline engine response to fuels applicable in all modes
- Coefficients $\{m_i, n_i, o_i, p_i\}$ and $\{q_i, r_i, s_i, t_i\}$ allow the engine response to fuel characteristics to vary in the RBEI and RCSI control modes, as additive corrections to the baseline fuel response.

The model form given in Eq. 6-1 is the form used to begin the process of finding the best-fit models for each operating mode; in that process, many of the terms are rejected for lack of statistical significance. For example, no statistically significant interactions were found between the engine control modes and the fourth fuel vector, so that coefficients p_i and t_i do not appear in any of the final models. More generally, the choice of terms for any dependent variable will vary across for the three operating modes, and the terms will differ among the dependent variables. There are 15 dependent variables measured in each three operating modes, so that the 45 different predictive models feed into the engine response models and simulators[1].

The predictive models were estimated using the Ordinary Least Squares (OLS) method as implemented by the SAS Proc REG and Proc GLM procedures. A stepwise selection process was used to search for the best-fit model for each dependent variable and operating mode. The selection process began with consideration of the full model form (Eq. 6-1) and used a backward

[1] Note that one of the 15 dependent variables is an engine output measure (either IMEP or dPdCA, depending on operating mode) that is used as an experiment target and is not free to vary across tests. It is not truly a dependent variable in the statistical sense.

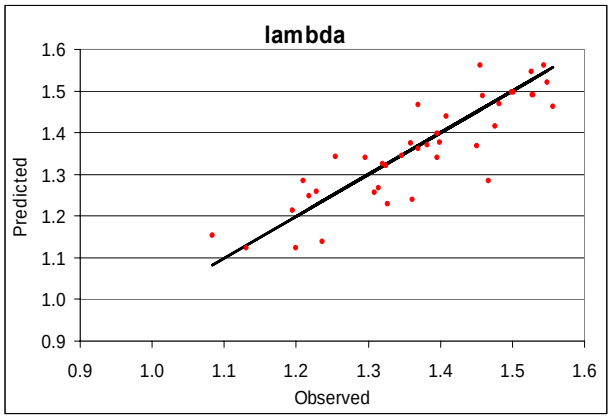
elimination process to reduce the model to only the statistically significant terms. The terms representing statistical controls for combustion phasing, engine output, and intake air temperature were required to be included in the model to assure that variation caused by experimental imprecision was fully controlled. Candidate terms were dropped from the model when they failed to reach at least the p=0.05 level of significance (95 percent confidence). Terms dropped in earlier stages could re-enter and remain in the model if they achieved the p=0.05 level of significance at a later stage. The resulting best-fit models achieve the highest R^2 using those terms from Eq. 6-1 that achieve the p=0.05 level of significance or better. The presence of a term in the model means that one can be 95 percent confident that the observed effect does not arise merely by random chance. The absence of a term means that the effect cannot be detected with the desired confidence; an effect of some size may be present, but too small to detect in this data, or the term may have no effect on the dependent variable.

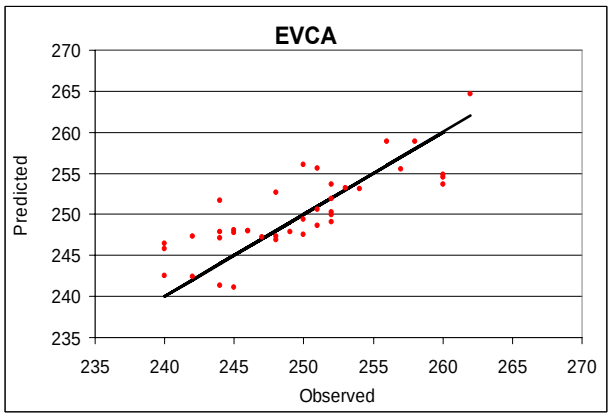
The experimental testing obtained repeated tests (normally five) for each fuel and control mode. The repeat tests have been reduced to an average test result for each fuel and control mode, and it is the averaged values that are used in estimating the predictive models. At the start of the analysis, consideration was given to the possibility that individual tests covered sufficient range in combustion timing and engine output that some information could be obtained on the variation of engine response similar to "sweeps" in the two variables. However, the engine testing was sufficiently well-controlled to target values that the deviations offer no more than localized noise in the observed values, and not information on performance across a useful range of combustion timings and engine outputs.

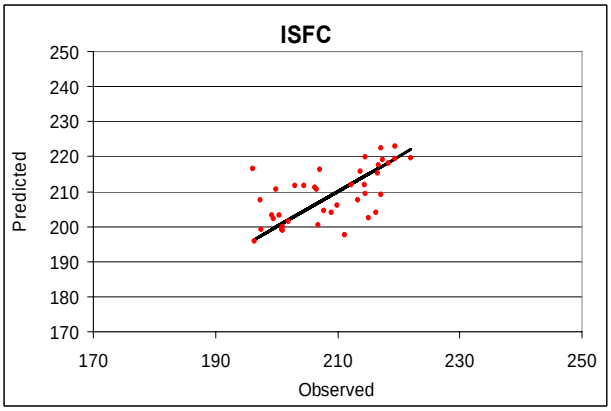
A subset of the entire engine test database was used to estimate the models. As previously described in Section 3, the test data associated with fuel B50D50 was struck from the analysis based on the evidence that the fuel characterizations were inconsistent or incomplete due errors in blending, measurement, or reporting. Further, the engine test data for indolene fuels were not used in estimating the predictive models, which are based solely on the experimental fuels. Further, an approach was adopted that controlled for the effects of experimental imprecision and engine drift over time that incorporates the statistical controls directly in the predictive models.

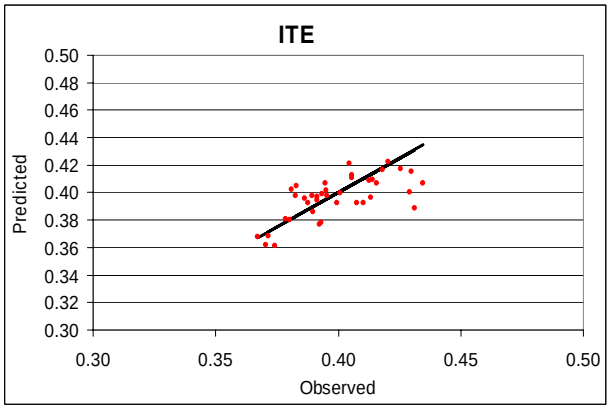
6.2.2 Predictive Ability of Base Engine Models

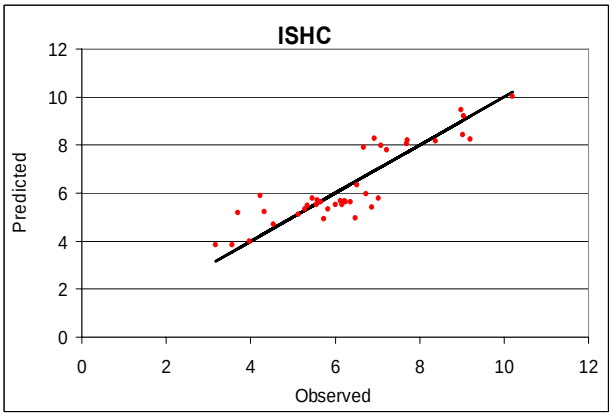
The quality of statistical models is perhaps most easily understood in the form of scatter plots of the observed and predicted values. Figure 6.1 presents such plots for the predictive models developed for Mode 2 (2000 rpm). Observed values are plotted on the horizontal axis, while the corresponding values predicted by the statistical models are plotted on the vertical axis. The solid black line marks the diagonal where the observed and predicted values are equal. In general, the predictive models do an acceptable job of representing the overall trend in engine performance observed in Mode 2, with a varying degree of scatter (unexplained variation) about the diagonal line depending on the variable. The plot for IMEP shows no scatter, because it was an experimental target in the Mode 2 testing and is not truly a dependent variable for the mode, while the plot for PCP shows the greatest scatter. One also sees that the models for CombEff and ExhT are backed by very little data, since the variables were measured only in Phase 2b (after engine rebuild). These models are not likely to be reliable no matter how good the fit to the data. In general, the predictive ability of the models is comparable to that obtained in many other successful studies in the area of motor vehicle performance.

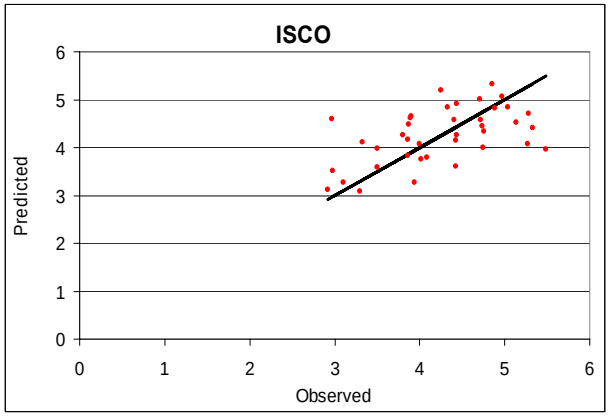

lambda
Predicted
Observed
0.9
1.0
1.1
1.2
1.3
1.4
1.5
1.6


EVCA
Predicted
Observed
235
240
245
250
255
260
265
270


ISFC
Predicted
Observed
170
190
210
230
250


ITE
Predicted
Observed
0.30
0.35
0.40
0.45
0.50


ISHC
Predicted
Observed
0
2
4
6
8
10
12


ISCO
Predicted
Observed
0
1
2
3
4
5
6

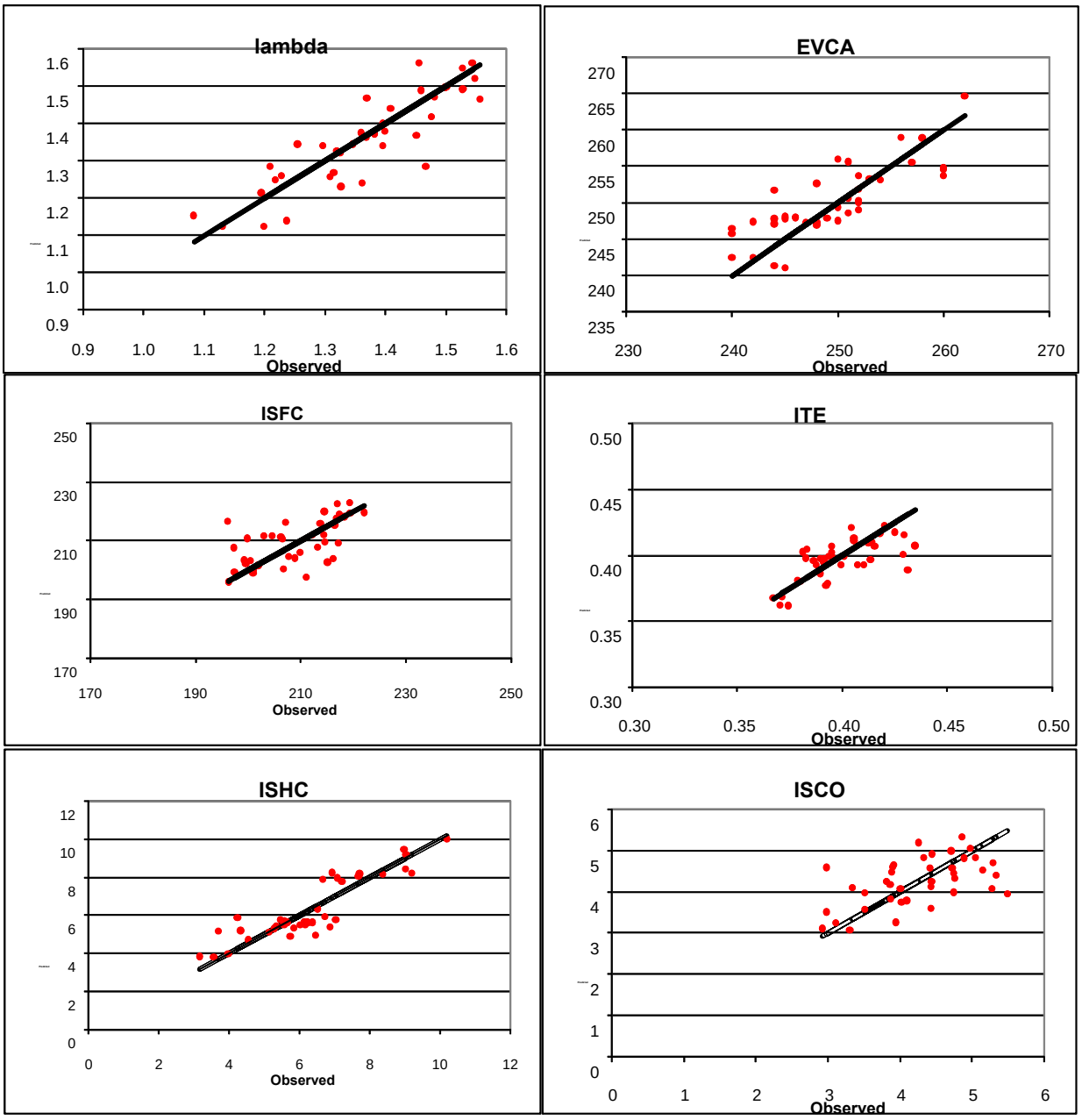


Figure 6.1: Comparison of Predicted vs. Observed Values for Mode 2 Engine Simulator Models.

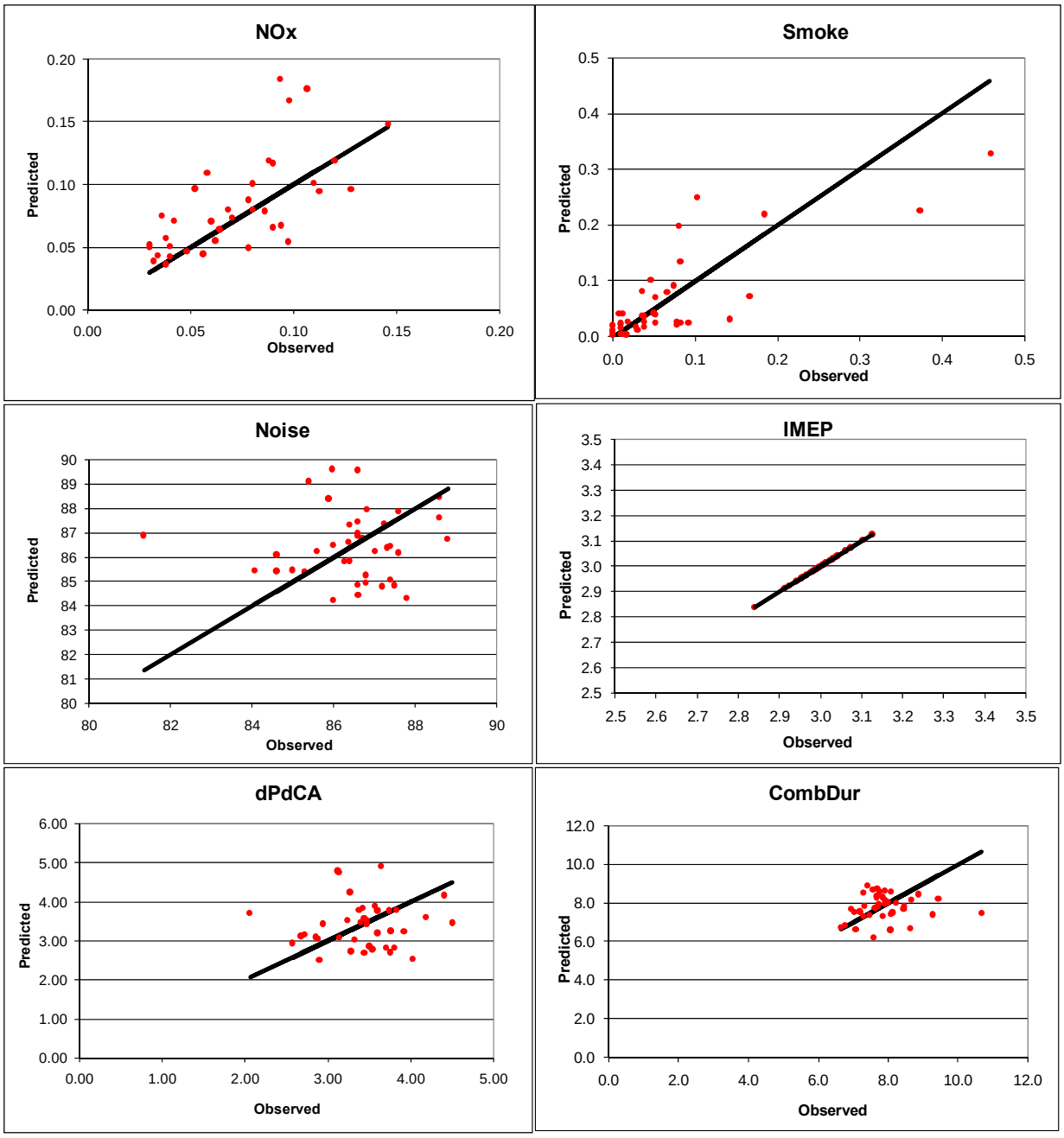


Figure 6.1 (continued)

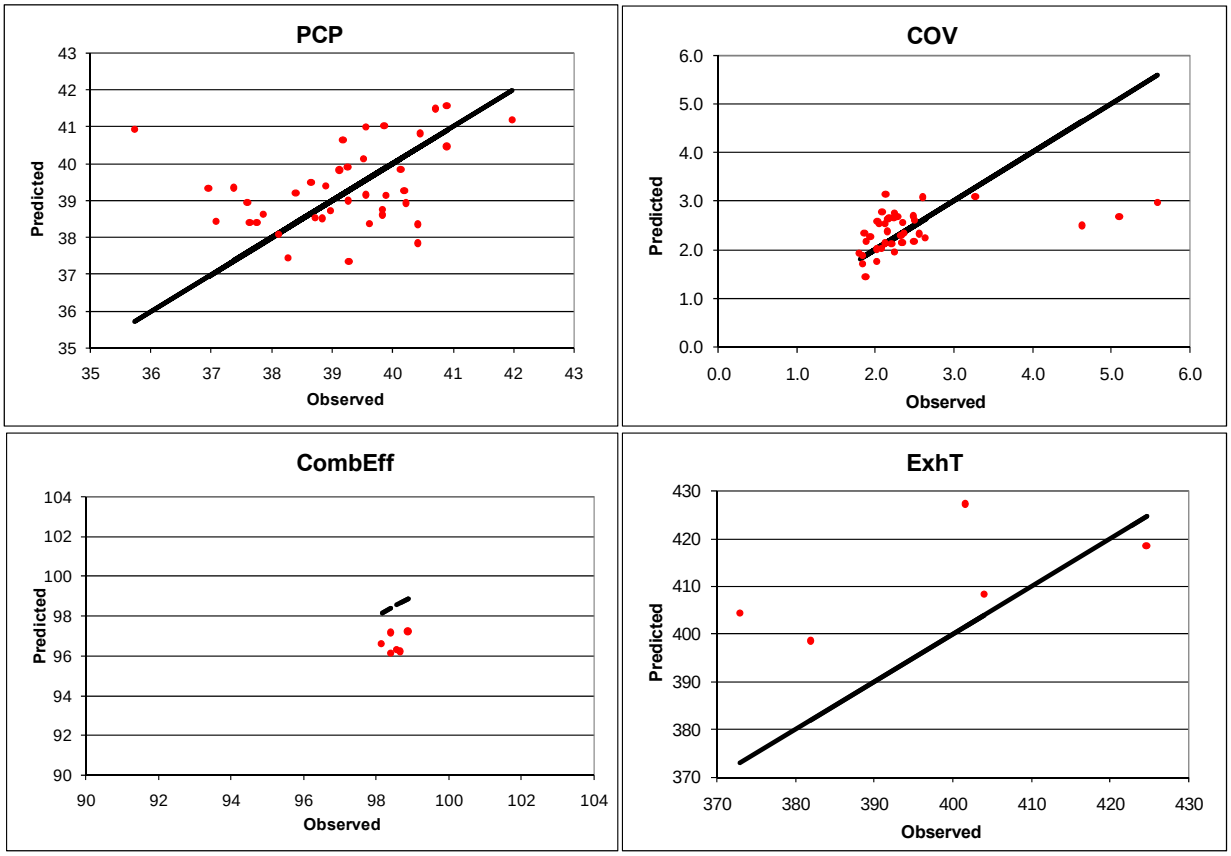


Figure 6.1 (continued)

The scatter plots display the total predictive power of the models with respect to the data, but do not indicate how the predictive power is distributed among the many terms of Eq. 6-1. In general, the predictive power of a model will be mixed among the statistical control variables and the variables representing the fuel and engine control effects that were the subject of the testing. A dependent variable may be strongly affected by the experimental imprecision in hitting the engine operation targets, it may be strongly affected by fuel characteristics and engine control modes, or both groups of variables may have strong effects.

Tables 6.2 through 6.4 summarize the predictive ability of the best-fit models in each operating mode in terms of the overall R^2 statistic and a qualitative classification of model quality based on the R2: *Strong* for models having $R^2 \geq 0.70$; *Good* for models having $R^2 \geq 0.50$; and *Weak/Poor* for models having $R^2 < 0.50$. The tables also offer insight on the question of the relative contributions from the statistical control (or adjustment variables) related to deviations from engine operating targets and the representations of test phase effects and fuel/engine control variables.

Using Mode 1 (Table 6.2) as an example, the best-fit model for lambda achieves an overall R^2 of 0.724, which is classified as Strong. However, the predictive ability is almost entirely the result of the contributions from the statistical control variables ($R^2 = 0.701$), while the fuel and engine

control variables add only 0.023 to the overall R^2. In Mode 1, the variance in lambda is primarily determined by changes in the engine operating conditions of the tests and by changes in engine condition across the phases. Fuel characteristics and the engine control mode explain only a very small part of the observed variance.

Looking across the tables, one can see that engine performance in Mode 1 (1000 rpm) is dominated the factors represented by statistical controls, while the effects of fuels and engine control mode are generally modest. Only EVCA and Smoke depend significantly on fuels and engine control (R^2 contribution $\geq$ 0.50), while ISCO shows a modest dependence (R^2 contribution $\geq$ 0.30). A majority of the predictive models are classified as Strong or Good overall; however, ISFC and ITE are poorly predicted in Mode 1, as is ISHC.

In Mode 2 (2000 rpm), engine performance continues to be strongly affected by the factors represented by the statistical controls, but it is now more strongly affected by fuels and engine control mode effects (Table 6.3). EVCA and Smoke are, again, significantly affected by fuels and engine controls (R^2 contribution $\geq$ 0.50), but six variables (ISHC, ISCO, ISNOx, CombDur, PCP, and COV) now show a modest dependence on those variables (R^2 contribution $\geq$ 0.30). All of the predictive models except for COV are classified as Strong or Good overall.

In Mode 3 (3000 rpm), engine performance is as strongly affected by fuels and engine control mode as it is by the factors represented by the statistical controls (Table 6.5). ISCO, ISNOx and COV are the variables most strongly affected by fuels and engine controls (R^2 contribution $\geq$ 0.50), while seven variables (lambda, EVCA, ISFC, ITE, ISHC, Smoke, and IMEP) show a modest dependence (R^2 contribution $\geq$ 0.30). All predictive models are classified as Strong, except for IMEP (Good) and CombDur (Weak/Poor).

In general, the quality of the engine response models increases when one moves from Mode 1 to Modes 2 and 3 with higher engine speeds and engine output. The role of fuels and the engine control mode also becomes more important, particularly in Mode 3 at high speed (3000 rpm) and with the test condition targeting a maximum rate of rise in cylinder pressure.

Table 6.2: Summary of R^2 Statistic for Mode 1 (1000 RPM) Engine Simulator Model.

Dependent Variable	Best Fit Model		R^2 Contributions From	
	R^2	Classification	Statistical Controls	Fuel and Engine Control Effects
Lambda	0.724	Strong	0.701	0.023
EVCA	0.697	Good	0.166	0.531
ISFC	0.269	Weak/Poor	0.136	0.133
ITE	0.309	Weak/Poor	0.170	0.138
ISHC	0.454	Weak/Poor	0.290	0.164
ISCO	0.833	Strong	0.504	0.329
ISNOx	Emissions barely above detection threshold			
Smoke	0.668	Good	0.114	0.554
Noise	0.553	Good	0.470	0.083
IMEP	Not a dependent variable			
dPdCA	0.782	Strong	0.608	0.174
CombDur	0.703	Strong	0.476	0.227
PCP	0.910	Good	0.838	0.072
COV	0.688	Strong	0.570	0.118
CombEff	0.992	Strong	0.969	0.023
ExhT	0.979	Strong	0.879	0.099

Table 6.3: Summary of R^2 Statistic for Mode 2 (2000 RPM) Engine Simulator Model.

Dependent Variable	Best Fit Model		R^2 Contributions From	
	R^2	Classification	Statistical Controls	Fuel and Engine Control Effects
Lambda	0.725	Strong	0.439	0.286
EVCA	0.571	Good	0.025	0.546
ISFC	0.565	Good	0.335	0.230
ITE	0.632	Good	0.390	0.242
ISHC	0.844	Strong	0.383	0.462
ISCO	0.793	Strong	0.424	0.369
ISNOx	0.855	Strong	0.355	0.499
Smoke	0.694	Good	0.093	0.601
Noise	0.720	Strong	0.495	0.225
IMEP	Not a dependent variable			
dPdCA	0.802	Strong	0.636	0.166
CombDur	0.694	Good	0.349	0.345
PCP	0.834	Good	0.509	0.326
COV	0.469	Weak/Poor	0.044	0.425
CombEff	0.978	Strong	0.685	0.293
ExhT	0.996	Strong	0.928	0.069

Table 6.4: Summary of R^2 Statistic for Mode 3 (3000 RPM) Engine Simulator Model.

Dependent Variable	Best Fit Model		R^2 Contributions From	
	R^2	Classification	Statistical Controls	Fuel and Engine Control Effects
Lambda	0.847	Strong	0.388	0.459
EVCA	0.865	Strong	0.488	0.377
ISFC	0.810	Strong	0.476	0.334
ITE	0.842	Strong	0.530	0.312
ISHC	0.890	Strong	0.394	0.496
ISCO	0.888	Strong	0.265	0.622
ISNOx	0.860	Strong	0.223	0.638
Smoke	0.826	Strong	0.420	0.406
Noise	0.705	Strong	0.537	0.167
IMEP	0.524	Good	0.116	0.408
dPdCA	Not a dependent variable			
CombDur	0.474	Weak/Poor	0.335	0.139
PCP	0.913	Strong	0.782	0.131
COV	0.825	Strong	0.166	0.660
CombEff	0.989	Strong	0.825	0.164
ExhT	0.996	Strong	0.981	0.015

6.3 Engine Models for Oxygenated Fuels

In Phase 2, oxygenated fuels were created by splash blending ethanol in varying amounts with five selected base fuels to create five series of oxygenated fuels (Table 6.5). Two fuel series were blended and tested in Phase 2a. Fuel A79D21 was blended at three ethanol levels (nominally E10, E20 and E30), with the measured ethanol content actually reaching 35% by volume for the E30 fuel. The C100 base fuel was blended to a single E20 level, but as described in Section 3.3, the C100E20 fuel was excluded from the analysis because its reported distillation curve was inconsistent with a 20 percent blend of ethanol with C100. Three fuel series were blended and tested in Phase 2b, using the base fuels B100, B77D23 and C50D50 blended with ethanol at two levels (nominally E15 and E30).

With exclusion of the C100E20 fuel, the dataset contains nine different oxygenated fuels, with measured ethanol contents ranging from 11% to 35% and created from four different base fuels. As will later be seen, the fact that only one fuel series was tested in Phase 2a introduces confounding between possible effects related to the characteristics of the single base fuel and known differences in the engine condition between Phases 2a and 2b due to the engine rebuild. In general, this is a relatively small dataset, and the analysis has been intentionally cautious to avoid over-interpreting the data.

Table 6.5: Summary of Oxygenated Fuel Design.

Base Fuel	Test Phase	Nominal Blending Levels (vol %)	Measured Ethanol Contents (vol %)	Number of Fuels
A79D21	2a	E10, E20, E30	11%, 22%, 35%	3 total
C100	2a	E20	21%	excluded
B100	2b	E15, E30	16%	2 total
B77D23	2b	E15, E30	16%, 29%	2 total
C50D50	2b	E15, E30	16%, 30%	2 total

6.3.1 Mathematical Formulation

The testing of oxygenated fuels during Phase 2 can be viewed as a sub-experiment that is distinct from the Phase 1 experiment designed to test engine response to base fuel characteristics. The key points of distinction are that the oxygenated fuels were splash blended with no effort to control RVP to the 7 psi level of the base fuels (or to control other fuel property changes) and that the testing was conducted late in the program (entirely in Phase 2 and predominantly in Phase 2b after the engine rebuild). After consideration of these distinctions, the decision was made to conduct the analysis of engine response to oxygenated fuels as a separate analysis using a paired variable technique, rather than as part of the analysis of engine response to base fuels.

For any given test, the paired variables are the engine response on the oxygenated fuel and the engine response on the base fuel.

To include the oxygenated fuels with the base fuels in a composite analysis of engine response would require specifying a regression model that accurately captures all terms that are germane to modeling how the engine responds to the base characteristics of fuels, to the oxygen content of the fuels, and to potential interactions among base and oxygenated fuels. In a properly specified model – meaning one that contains all necessary terms – and with suitable data, a conventional regression analysis should be able to separate the effects of each term and provide unbiased estimates of the response coefficients. However, the process depends critically on the proper specification of the regression model. In cases where (as is true here) the data do not constitute an orthogonal experiment, the regression fit can lead to unsatisfactory results due to aliasing – including the mixing of effects between base fuel characteristics and oxygen content – whenever needed terms are omitted or given mathematical forms that depart from the physical reality. Such concerns are always present to some degree in the analysis of non-orthogonal data, but were judged to be of greater concern here because of the relatively small number of oxygenated tests.

In a paired analysis, the engine performance on the i^{th} oxygenated fuel is compared to the engine performance on the corresponding base fuel from which the i^{th} fuel was blended. The difference in performance between the oxygenated and base fuels defines the engine response to fuel oxygenation; statistical models are then built to explain the response as a function of ethanol content, engine control mode, and possibly base fuel characteristics and other variables. The primary advantages of this approach are:

- Calculation of the differential engine response gives a direct and unambiguous measure of how fuel oxygenation affects engine performance. In a combined analysis involving both base and oxygenated fuels, the engine response to fuel oxygenation would be inferred (rather than calculated) by segregating the terms involving oxygen content from those involving other effects.
- Estimation of the oxygen effect is not dependent on the quality of the statistical model for the base fuel effects, because the base fuel effect is present in both of the paired tests (the oxygenated fuel and its corresponding base fuel).

The primary drawback to the paired analysis is that it is performed separately from the analysis of base fuels and its results must be grafted onto the results of the latter in order to make predictions for all fuels. This has been done in a seamless way in the engine simulators created for this study and should not be apparent to users.

The dependent variables in the paired analysis are the calculated differences in engine response, taking into the account the use of a logarithmic transform for some variables (as documented in Table 6.1). When the engine response variable is transformed to a logarithm in the analysis of base fuel effects, the dependent variable in the paired analysis is computed as the difference in logarithms between the oxygenated and base fuel tests; this is equivalent to the logarithm of the ratio of the measured values between the oxygenated and base fuel tests. In other cases, the dependent variable is computed as the difference in measured values between the oxygenated and base fuel tests.

Using the notation $P_{i,j}$ to denote the i^{th} performance metric P_i (transformed or not) for the j-th engine test, the mathematical form of the engine response model for any operating mode can be written as follows:

$$\Delta P_{i,j} = P_{i,j} - P_{i,Base} \qquad \text{(Eq. 6-2)}$$

where $P_{i,Base}$ refers to the corresponding engine performance metric for the base fuel. Then, the regression models in the paired analysis take the form:

$$\begin{aligned}\Delta P_{i,j} = {} & a_i + b_i * EtOH + c_i * EtOH^2 \\ & + d_i * dRCEI * EtOH \\ & + e_i * [(MFB50_j - MFB50_{target}) - (MFB50_{j,Base} - MFB50_{target})] \\ & + f_i * [(EngOutput_j - EngOutput_{target}) - (EngOutput_{j,Base} - EngOutput_{target})] \\ & + g_i * [(IntakeT_j - IntakeT_{target}) - (IntakeT_{j,Base} - IntakeT_{target})]\end{aligned} \qquad \text{(Eq. 6-3)}$$

The coefficients b_i and c_i measure the engine response to ethanol content (measured in vol %), with the presence of a quadratic coefficient allowing for a change in the response at higher ethanol levels (i.e., accelerating or diminishing returns). The coefficient d_i represents a potentially different engine response to ethanol in the RCEI mode. Phase 2 did not conduct tests in the RCSI mode. The dummy variable RBEI is multiplied by the EtOH content of the fuel to account mathematically for the requirement that the differential engine response to an oxygenated fuel must go to zero as the ethanol content goes to zero. The coefficients { e_i, f_i, and g_i } represent the statistical controls for the deviation of experimental conditions from the nominal values of the targets for each operating mode. These terms allow for the fact that the tests on the oxygenated fuel and on the base fuel will deviate in different ways from the targets.

The data used in the analysis are the repeat-test weighted averages for the nine oxygenated fuels and the corresponding four base fuels, giving 13 data points in each engine control model (RCEI and RBEI) and 26 data points across all fuels. The test data for base fuels are introduced with computed values of zero for the values of the dependent variables, but an intercept is fit by the regression to recognize that the observed values for base fuels do not necessarily lie on the response line due to random variation. The response lines for the paired variables are not forced to pass through the origin (zero response at zero oxygen content), but all of the lines do so within the standard errors of estimate.

Because of the small number of data points, it is unlikely that all of the terms in Eq. 6-3 can be reliably estimated as a starting point for stepwise regression. Therefore, a forward selection technique was employed in which the starting model consisted of the three statistical control terms and the linear term in EtOH. The statistical control terms were required to be included in the models to assure that experimental imprecision was fully controlled, but the EtOH and other terms given in Eq. 6-3 were included only when they achieved the p=0.05 level of significance or better.

6.3.2 Predictive Ability of Oxygenated Fuels Models

Tables 6.6 through 6.8 summarize the statistical models for engine response to oxygenated fuels by operating mode. Each table gives the R^2 statistics for the dependent variables measured in the mode, along with characterization of the EtOH effect as linear or quadratic and the estimated size of the EtOH effect computed at 15% by volume. The instances where the effect differs by engine control mode are also given, as are the percentage difference compared to the baseline RCEI control mode. For example, the analysis indicates that ISFC is 15% higher under the RBEI control mode when the engine is operated at 1000 rpm.

In most instances, the best-fit statistical models indicate that no statistically significant effect of fuel oxygenation can be detected, given the number of data points and the variance unavoidably introduced by experimental imprecision. Where an effect of fuel oxygenation is detected, the effect of fuel oxygenation is generally linear with respect to the blending volume; in the two instances where pure quadratic effects were found, these show only small degrees of curvature and would be well-approximated by linear slopes. In only six instances overall is the effect of fuel oxygenation found to depend upon engine control mode, and in each case the result is to say that the effect is present in one mode but not the other.

A substantial degree of caution must be exercised in considering these results because of the small number of tests. For example, detecting a fuel oxygen effect in one control mode, but not the other, is fully consistent with the hypothesis that an effect is present and of the same size in both control modes, but can be detected in only one mode because of the small sample size. Further, the detection of an effect for a particular dependent variable in one operating mode, but not in others, is difficult to interpret conclusively. It is certainly reasonable to find that the oxygen effect appears in different metrics in different modes because of the substantially different engine speed, fueling rates and cylinder temperature/pressure profiles by mode. On the other hand, there are 45 different models (dependent variables times operating modes) and multiple statistical tests of significance at the $p=0.05$ level in each case. Given the large number of tests, it is likely that some of the findings, at least, occur because of random chance. Across operating modes, the data indicate that fuel oxygenation has generally small effects on a small number of performance metrics and that there is little evidence that the effects are of substantially different size between the RBEI and RCEI control modes. A later section showing the results of the parametric studies will indicate that the primary effects of fuel oxygenation appear to be related to dilution of the energy density of the bulk fuel and effects, both adverse and beneficial, related to reduction in cylinder temperatures.

Table 6.6: Summary of Engine Response Model for Oxygenated Fuels: Mode 1 (1000 rpm).

	R^2	Type of EtOH Effect	Size of Effect at 15% EtOH	Dependence on Control Mode (versus RCEI)
Lambda	0.68	–	–	–
EVCA	0.01	–	–	–
ISFC	0.57	–	–	+15% for RBEI
ITE	0.38	–	–	–
ISHC	0.24	–	–	–
ISCO	0.87	Linear	+14%	–
ISNOx	n/a	–	–	–
Smoke	0.40	–	–	–
Noise	0.42	–	–	–
IMEP	Not a dependent variable			
dPdCA	0.93	–	–	–
CombDur	0.74	Linear	+7%	–
PCP	0.85	–	–	-7% for RBEI
COV	0.71	–	–	–
CombEff	0.78	Linear	-1%	–
ExhT	0.46	Linear	-5°C	–

Table 6.7: Summary of Engine Response Model for Oxygenated Fuels: Mode 2 (2000 rpm).

Dependent Variable	R^2	Type of EtOH Effect	Size of Effect at 15% EtOH	Dependence on Control Mode (versus RCEI)
Lambda	0.30	Linear	+5%	–
EVCA	0.30	–	–	–
ISFC	0.80	–	–	+15% for RBEI
ITE	0.38	–	–	–
ISHC	0.44	Linear	+9%	–
ISCO	0.29	–	–	+50% for RBEI
ISNOx	0.37	–	–	–
Smoke	0.43	–	–	–
Noise	0.71	–	–	–
IMEP	Not a dependent variable			
dPdCA	0.67	–	–	–
CombDur	0.59	–	–	–
PCP	0.69	Linear	-2%	–
COV	0.28	–	–	+91% for RBEI
CombEff	0.09	–	–	–
ExhT	0.34	–	–	–

Table 6.8: Summary of Engine Response Model for Oxygenated Fuels: Mode 3 (3000 rpm).

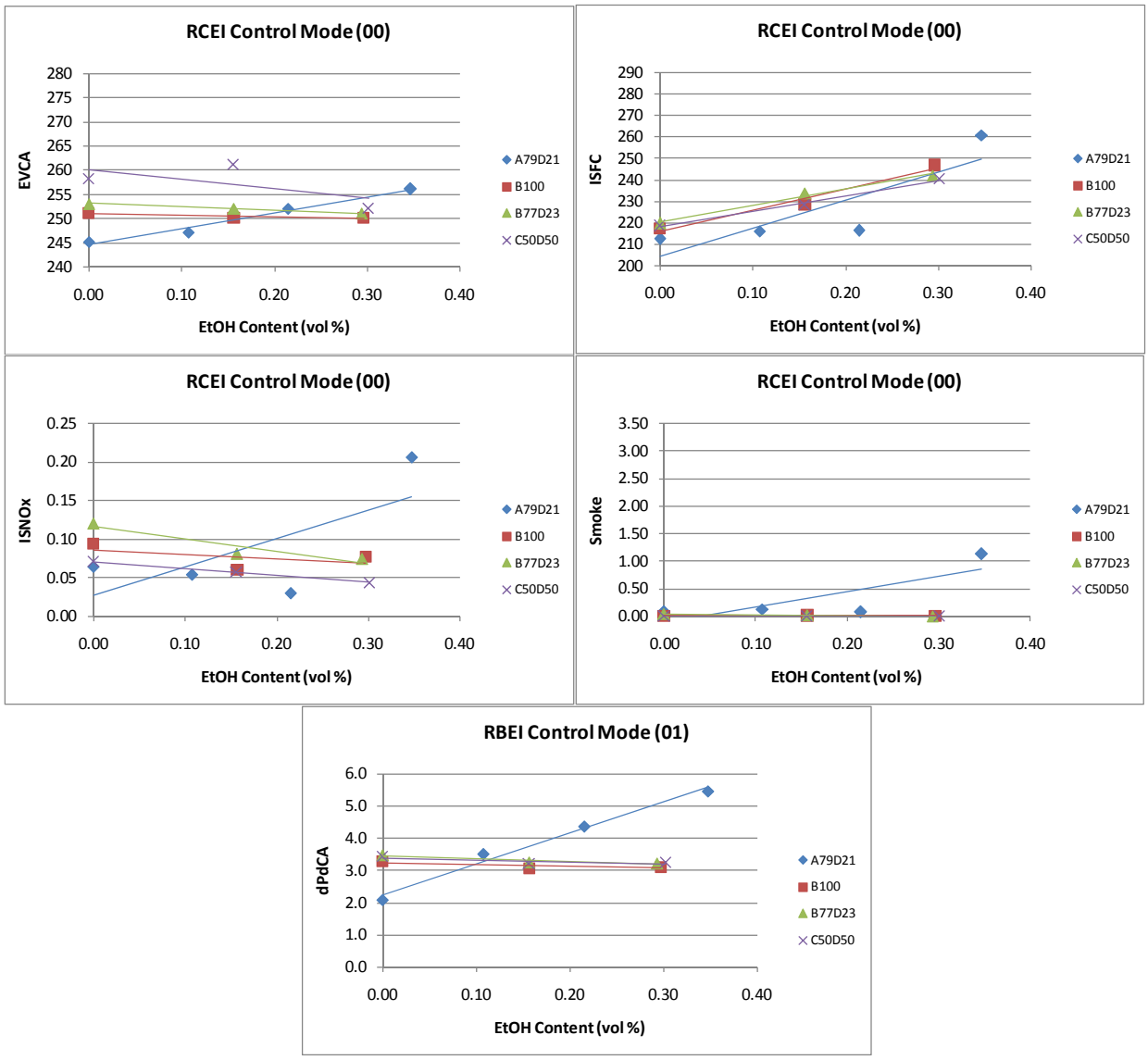
Dependent Variable	R^2	Type of EtOH Effect	Size of Effect at 15% EtOH	Dependence on Control Mode (versus RCEI)
Lambda	0.40	–	–	–
EVCA	0.37	–	–	–
ISFC	0.75	Quadratic	+3%	–
ITE	0.06	–	–	–
ISHC	0.42	–	–	–
ISCO	0.40	–	–	–
ISNOx	0.59	Linear	-27%	–
Smoke	0.19	Linear	126%	–
Noise	0.82	Linear	-0.2	–
IMEP	0.39	–	–	–
dPdCA	Not a dependent variable			
CombDur	0.35	–	–	–
PCP	0.61	Quadratic	-2%	0% for RBEI
COV	0.36	–	–	–
CombEff	0.66	–	–	–
ExhT	0.88	Linear	-18°C	–

Figure 6.2 emphasizes the need for caution in interpreting these results through the divergent trends and scatter seen in the individual graphs making up the figure. This figure displays the Mode 2 (2000 rpm) test data for five important metrics of engine performance as a function of fuel ethanol content. Only the RCEI control mode results are shown for clarity. While there is scatter among the tests and some apparent differences among the base fuel series, the graphs suggest overall that increasing oxygen content does not affect the EVCA needed to achieve combustion phasing and engine output targets. ISFC is increased, ISNOx may or may not be affected, while Smoke is unaffected, as is dP/dCA.

In particular, note how the graphs indicate that the A79D21 fuel series (the only fuel series run in Phase 2a) often shows a different trend than the three fuel series run in Phase 2b (following the engine rebuild), which are generally more consistent with each other. Base fuel A79D21 has the highest score for the Vector 1 characteristic among the four base fuels that were oxygenated, while the fuels tested in Phase 2b were either moderate or low in that characteristic. It is also known that the engine rebuild affected performance, particularly in making the engine less susceptible to producing Smoke. The dilemma is how to interpret the divergent results for the A79D21 fuel series, since for some plots like ISFC, ISNOx, and Smoke, the data for A79D21E30 diverge from the trends for the other base fuels (see Figure 6.2). Unfortunately, we do not have enough date to determine if this is a base fuel effect, from the engine rebuild between Phases 2a and 2b, or from some experimental measurement problem.

It is not possible to answer these questions conclusively given the available data on oxygenated fuels. The statistical analysis is not equipped to distinguish between effects related to base fuel characteristics versus differences between the test phases induced by the engine rebuild. Only one fuel series was tested in Phase 2a, and its base fuel is the one most different from the other fuels. This confounding of base fuel and test Phase effects can be resolved only with further testing.

Figure 6.2: Ethanol Effects on Engine Performance: Experimental Data, Mode 2 (2000 rpm).



7. USE OF MODELS

Once the models have been developed and verified, they can be used to conduct various parametric studies to show trends uncovered in the data. These studies can encompass engine control mode, engine operating mode, or fuel variable effects. For these types of studies, the variable or variables of interest would be set up in a parametric manner to span the range of the variables of interest, within the range that the experiments encompassed. Each variable is normally set to incremental levels, and the number of members in a parametric run is the multiplicand of the number of levels for each variable. The other variables in the problem can be set to a nominal value, an optimum value, or allowed to vary over their range in the experimental data. When the models are exercised in this manner, a given output variable may or may not

show a correlation to a given input variable and this correlation is used to study and demonstrate the effects being studied. Examples of questions which will be studied in this manner are:

- What engine control mode provides the best overall engine performance when evaluated using a mid-point fuel?
- Are there systematic variations in engine response for the different test phases?
- What engine variables show correlation to fuel changes?
- What is response of engine to changes in sensitivity or chemistry at constant (R+M)/2?

Using the models to perform parametric studies in this manner provides several advantages. First, it allows all the data to participate in studying the trends, since all the data were used to construct the models. Second, since these models are easy to use and will be delivered to CRC as part of this project, other studies can be conducted by members, rather than being limited to studies and conclusions included in this report. This means that this report can be considered a launch point and instruction manual for the use of these models by others to continue the study of fuel and engine effects. In these parametric studies, the fuels are represented by principal components and the resulting vector values for each parametric fuel. Using a vector representation of the fuels provides the ability for all the fuel variables to participate in the model, while retaining the correlations between fuel variables that exist in the original fuels. The use of principal components does not eliminate correlation between fuel variables; rather it moves the need to choose between correlated variables to a point after the construction of the models and the parametric studies. The authors believe that the shift of the decision point to the final stage of the analysis provides the user with the most options for studying engine and fuel responses.

8. PARAMETRIC STUDIES

This section describes a number of parametric studies which were performed using the fuels and engine models. The studies were selected to demonstrate the use and capabilities of the modeling tools and also to point out interesting trends found in the data. There are obviously a very large number of other studies which could be performed using the fuels and engine simulators supplied with this report. In all cases, one must define inputs to fall within the range of experimental data and interpret the results for likelihood and ability to be explained.

8.1 Fuel Selection for Parametric Studies

The following studies were done with either a single fuel of ‘average’ properties and chemistry, with a general parametric fuel set of 249 members which was constructed for this purpose, or with a set of parametric fuels screened to have a fixed (R+M)/2 and as wide a range of chemistry and sensitivity as possible. The single fuel of average properties was constructed by setting all PCA vector values to zero, in order to determine the midpoint or average fuel for the experimental fuels. This can be done with the fuel panel simulator by using the “Vectors” tab (described in Section 9.1.3) and setting the values of all of the vectors to zero. This results in a fuel with properties and chemistry as shown in Table 8.1.

Table 8.1: Average Fuel (Vector Scores = 0.0).

PROPERTIES, CHEMISTRIES	UNITS	VALUE
n-paraffins except C4	%	9
n-paraffins, C4	%	4
iso-paraffins	%	43
cyclo-paraffins	%	10
olefins	%	9
aromatics	%	25
RVP	psi	7.1
RON	number	82.6
MON	number	77.5
sensitivity	number	5.1
specific gravity	number	0.740
T10	deg.F	161
T50	deg.F	234
T90	deg.F	321

A second, more general, parametric set of fuels was constructed by sampling the vector values systematically using three to five discrete steps over a range from minus one standard deviation to plus one standard deviation in a nested manner that itemizes all combinations of vector values. Fuel chemistry and properties were then estimated from the vector representation, and fuels that fell outside of the experimental range for chemistry and properties were eliminated. This created a cluster of 249 fuels which covered the central range of the experimental fuels. The coverage of these fuels in vector space, compared to the experimental fuels, is shown in Figure 8.1. The parametric fuels form a cluster which falls within the vector ranges of the experimental fuels and also cover the range of experimental properties and chemistry exhibited in the experimental fuels. Refer to Figure 4.1 for a larger version of these plots with only the experimental data points shown.

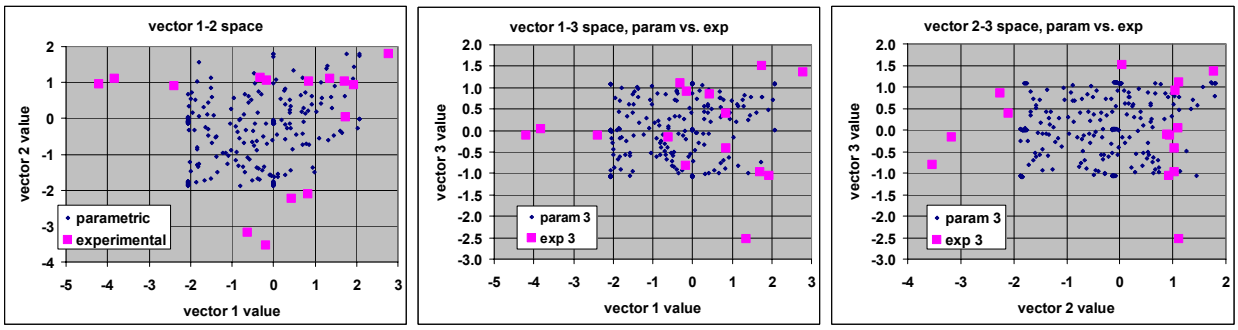


Figure 8.1: Coverage of Experimental and Parametric Fuels in V1, V2, and V3 Vector Space.

Finally, a set of 31 fuels with constant (R+M)/2 were constructed by first creating a set of 5000 fuels which randomly and evenly covered the range of vector values for the experimental fuels. The 5000 fuels were pared down to fit within the boundaries of the experimental vector pairs (leaving 1290 fuels). Then RON, MON, and (R+M)/2 were calculated using formulas extracted from the simulators, and 32 fuels were selected which fell between 86.9 and 87.1 (R+M)/2. These fuels were then passed through the fuel panel to calculate properties and chemistry for input to the engine simulators. The multiple math steps of this process resulted in a slightly larger deviation for (R+M)/2 than desired, but final fuels still fell within a band of ±0.5. This set of fuels is shown in Table 8.2. For these fuels, octane is developed by a combination of iso-paraffins, olefins, and aromatics. Fuels with large percentages of iso-paraffins have low sensitivity, and fuels with large percentages of aromatics or olefins have high sensitivity. At mid-sensitivity (≈5), aromatics and olefins can be traded against each other for constant sensitivity and varying chemistry.

Table 8.2: Parametric Fuels Selected for Constant (R+M)/2.

V1 value	V2 value	V3 value	nPxc4	nPc4	iP	cP	olefin	aromatic	RVP	RON	MON	sensitivity	(R+M)/2	SG	T10	T50	T90
-2.12	0.95	-1.01	0.06	0.00	0.29	0.07	0.29	0.28	7.1	93.0	82.0	10.9	87.5	0.747	148	213	283
-1.49	1.03	-1.25	0.07	0.01	0.29	0.06	0.23	0.33	7.1	93.0	82.2	10.8	87.6	0.754	153	222	291
-0.27	1.20	-1.67	0.09	0.01	0.28	0.06	0.14	0.43	7.1	92.9	82.1	10.8	87.5	0.772	165	241	305
-1.25	1.07	-1.31	0.08	0.01	0.29	0.07	0.22	0.34	7.1	92.5	81.7	10.8	87.1	0.757	155	225	293
-0.69	1.16	-1.50	0.09	0.01	0.28	0.06	0.17	0.39	7.1	92.4	81.6	10.8	87.0	0.766	161	235	300
-1.55	1.02	-1.22	0.07	0.01	0.29	0.07	0.24	0.32	7.1	92.6	81.9	10.7	87.3	0.752	152	220	290
-1.52	1.04	-1.21	0.07	0.01	0.30	0.07	0.23	0.32	7.1	92.4	81.8	10.6	87.1	0.753	153	222	291
-0.11	1.24	-1.70	0.09	0.01	0.28	0.06	0.12	0.44	7.1	92.5	81.8	10.6	87.2	0.774	166	244	307
-1.07	-1.03	-0.60	0.04	0.03	0.57	0.05	0.12	0.20	7.1	89.9	85.5	5.4	87.7	0.724	153	215	314
-0.31	-0.92	-0.86	0.05	0.03	0.55	0.04	0.06	0.27	7.1	90.1	84.7	5.4	87.4	0.736	160	227	323
-1.89	-1.14	-0.31	0.03	0.02	0.59	0.05	0.18	0.13	7.1	89.8	84.4	5.4	87.1	0.712	145	202	304
1.43	-0.67	-1.47	0.07	0.04	0.51	0.03	0.00	0.36	7.1	90.2	84.8	5.4	87.5	0.761	176	254	343
-0.33	-0.92	-0.85	0.05	0.03	0.55	0.04	0.06	0.27	7.1	90.1	84.7	5.4	87.4	0.736	160	227	323
-1.07	-1.02	-0.59	0.04	0.03	0.57	0.05	0.12	0.20	7.1	89.9	84.5	5.4	87.2	0.724	153	215	314
1.22	-0.70	-1.39	0.07	0.04	0.51	0.03	0.00	0.35	7.1	90.0	84.6	5.4	87.3	0.758	174	250	341
0.04	-0.87	-0.99	0.05	0.03	0.55	0.04	0.05	0.29	7.1	90.1	84.8	5.3	87.5	0.740	163	232	327
1.46	-0.66	-1.47	0.07	0.04	0.51	0.03	0.00	0.36	7.1	90.0	84.7	5.3	87.4	0.762	177	255	344
-0.23	-0.90	-0.88	0.05	0.03	0.55	0.04	0.06	0.27	7.1	89.9	84.6	5.3	87.3	0.737	161	229	324
1.45	-0.66	-1.47	0.07	0.04	0.51	0.03	0.00	0.36	7.1	90.0	84.7	5.3	87.4	0.762	177	255	344
-0.75	-0.97	-0.69	0.04	0.03	0.57	0.04	0.09	0.23	7.1	90.0	84.6	5.3	87.3	0.728	155	219	317
-2.01	-1.15	-0.25	0.02	0.02	0.60	0.05	0.18	0.11	7.1	90.1	84.8	5.3	87.5	0.710	144	200	304
0.01	-0.86	-0.96	0.05	0.03	0.55	0.04	0.05	0.29	7.1	90.1	84.8	5.3	87.5	0.740	163	232	327
1.26	-0.69	-1.40	0.07	0.04	0.51	0.03	0.00	0.35	7.1	89.8	84.5	5.3	87.2	0.759	175	252	342
-1.39	-1.06	-0.46	0.03	0.03	0.59	0.05	0.14	0.17	7.1	90.0	84.7	5.3	87.4	0.719	150	210	311
1.08	-0.71	-1.32	0.07	0.04	0.51	0.04	0.01	0.34	7.1	89.5	84.3	5.3	86.9	0.756	173	249	339
-1.49	-1.08	-0.44	0.03	0.03	0.59	0.05	0.14	0.16	7.1	89.6	84.4	5.2	87.0	0.717	149	208	309
-2.11	-1.17	-0.22	0.02	0.02	0.61	0.06	0.19	0.10	7.1	89.6	84.5	5.2	87.1	0.708	143	199	303
-0.05	-0.87	-0.94	0.05	0.03	0.55	0.04	0.05	0.27	7.1	89.7	84.4	5.2	87.1	0.738	162	230	326
-0.12	-0.88	-0.91	0.05	0.03	0.55	0.04	0.05	0.27	7.1	89.7	84.4	5.2	87.1	0.738	162	230	326
-0.14	-0.88	-0.90	0.05	0.03	0.55	0.04	0.05	0.27	7.1	89.7	84.4	5.2	87.1	0.738	162	230	326
-0.87	-0.98	-0.64	0.04	0.03	0.57	0.05	0.10	0.21	7.1	89.5	84.3	5.2	86.9	0.726	154	217	316
0.08	-2.98	-0.20	0.02	0.06	0.85	0.03	0.00	0.05	7.1	87.1	87.3	-0.1	87.2	0.705	159	219	346
-0.38	-3.03	-0.02	0.01	0.05	0.85	0.03	0.02	0.03	7.1	86.9	86.9	-0.1	86.9	0.698	154	212	340
0.20	-2.96	-0.24	0.02	0.06	0.85	0.03	0.00	0.05	7.1	86.9	87.1	-0.2	87.0	0.706	160	221	348

When plotted in a similar manner to Figure 8.1, as vector pairs, the graphs in Figure 8.2 show how the fuels were reduced to clusters which fall within the total range of experimental fuels. Interestingly, areas of fuels are also carved out inside of the outer boundaries. This is because all

but one of the experimental fuels were two component blends with D100 used to adjust octane values. Any fuel within the outer vector boundaries could have been blended with the provided components, but three and four component blends were not included as part of the experimental design. To the extent that one believes these fuels would blend linearly with the actual test fuels, the models can be used to predict their behavior. However, in the extreme, this logic could be extended to only testing the four blend components and calculating everything else, and this would not discover any non-linear blending or engine performance effects. These vector pair graphs in Figure 8.2 show the actual test fuels in red, the 1290 fuels selected from the 5000 fuels to fall within the vector, property, and chemistry boundaries in blue, and the 31 fuels which were selected for constant (R+M)/2 of 87 ± 0.1 in yellow and which are listed in table 8.2 above.

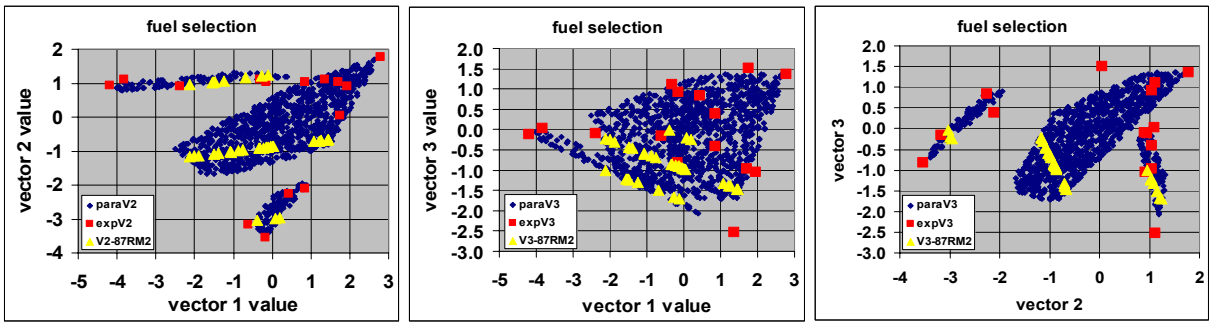


Figure 8.2: Vector Ranges for Fuels Selected to Fall Within Boundaries of Experimental Vectors and Further Selected for 87 (R+M)/2.

The three fuels or fuel sets discussed above will be used in the parametric studies which follow this section. It should be obvious that any number of parametric fuel sets can be constructed by systematic or random vector sampling or other means. Some fuel sets can be constructed fairly easily, and others, like the constant (R+M)/2 fuels, take a lot of hand work with the current, general modeling tools. No further automation has been applied to the process, because this would only be worthwhile if a given type of study were repeated multiple times. Any new fuel set developed should also be checked to determine that it falls within the range of experimental fuels for vector values, properties, and chemistry and then it can be used as inputs for the engine simulators. By using vector values and the fuel panel to originate any parametric study fuels, all available fuel variables are defined, are available for subsequent analysis, and out of range values are also flagged.

8.2 Control Mode Effects

During this research, the engine was operated at three operating points using three control modes. These modes are described elsewhere, and are named recompression early injection (RCEI), rebreathing early injection (RBEI), and recompression split injection (RCSI). These control modes were chosen because all were possible using the AVL engine and CRC wished to determine if any of the modes responded differently to fuel changes. In the parametric study described here, the fuel was set to the single mid-point fuel and test phase effects were zeroed out by setting the test phase input to an average value. Table 8.3 indicates the directional change in engine performance between the various control modes for each operating point. Plus means that performance improved, zero means that there was no change, and minus means that there was a negative change in performance for the given variable. Generally RCEI produced the most consistent performance with no negatives. This quick study indicates that RCEI is a good control mode to focus on for other parametric studies. It also indicates that a given control mode might be selected to improve one engine response at the expense of another.

Table 8.3: Directional Changes in Engine Performance Resulting from Control Mode Changes.

	1000 rpm			2000 rpm			3000 rpm		
variable	RCEI	RBEI	RCSI	RCEI	RBEI	RCSI	RCEI	RBEI	RCSI
ISFC	zero	zero	minus	zero	zero	zero	zero	zero	zero
ISHC	zero	zero	zero	zero	minus	zero	zero	minus	zero
ISCO	zero	zero	zero	zero	zero	zero	zero	minus	zero
ISNOx	zero	zero	zero	zero	zero	minus	zero	plus	zero
smoke	zero	zero	zero	zero	plus	zero	zero	minus	zero
noise	zero	zero	zero	zero	plus	zero	zero	minus	zero
dP/dCA	zero	zero	minus	zero	plus	zero	zero	zero	zero
PCP	zero	zero	zero	zero	plus	zero	zero	minus	zero
COV	zero	plus	zero	zero	minus	zero	plus	zero	zero

plus = BETTER, zero = SAME, minus = WORSE

8.3 Test Phase Effects

This test program took place over a three year period and included many changes during the three phases of the experiments (Phase 1, 2a, and 2b). These changes included a major engine rebuild, minor changes in intake and exhaust plumbing, a change in type of cylinder pressure transducer, engine wear, seasonal weather variations, ability to hit control targets, and a change in the project test engineer. Since fuels are being compared across the various test phases, it is important to know if there were any systematic offsets between the phases. In the engine simulator models, variations were compensated for by including dummy variables for the test phases (00 = Phase 1, 01 = Phase 2a, and 11 = Phase 2b), and including the variables for deviation from load or pressure rise rate set point, deviation from MFB50, and deviation from intake temperature set point. These variables were selected from the experimental data by correlating residuals of the engine simulator models without these adjustments and picking variables which correlated to the residuals and therefore could be assigned variation in more complex models. Figure 8.3 shows engine simulator outputs by phase and control mode for selected engine output variables. These outputs were calculated using the average fuel and are

shown only for the 2000 rpm model, with similar behavior exhibited for the other speeds. It is readily apparent from these graphs that engine performance is very different for phase 2b. Lambda is higher and EVCA is later, indicating easier ignition. This may be due to an inadvertent compression ratio change or due to better (new) piston rings and less compression losses. ISFC is worse, NO_x is higher, and HC, CO, and smoke are lower. Combustion duration is longer, which may be due to the higher lambda, lower exhaust residual operation or may be due to the change of cylinder pressure transducer from Kistler to AVL. The other observation from these graphs is that the differences in engine performance between control modes are similar across all the test phases. This is because of the study's focus on modeling the entire dataset, rather than concentrating on small regions of the data in order to identify large, more universal data trends. Studies like this can also identify desirable features of the control modes: for example RBEI exhibits lower smoke, noise, dP/dCA, and PCP and higher HC, CO, and COV. This indicates that control modes might be selected to optimize a particular operating condition or solve a particular operating problem. Note that RCSI was not run in Phase 2a or 2b, but that the model allows prediction of results by applying the same response measured in Phase 1. Further studies could be done using the supplied models, and additional experiments could be run to study control modes or operating modes in more depth. In general, RCSI control mode produced very similar results to RCEI when interpreted through the model.

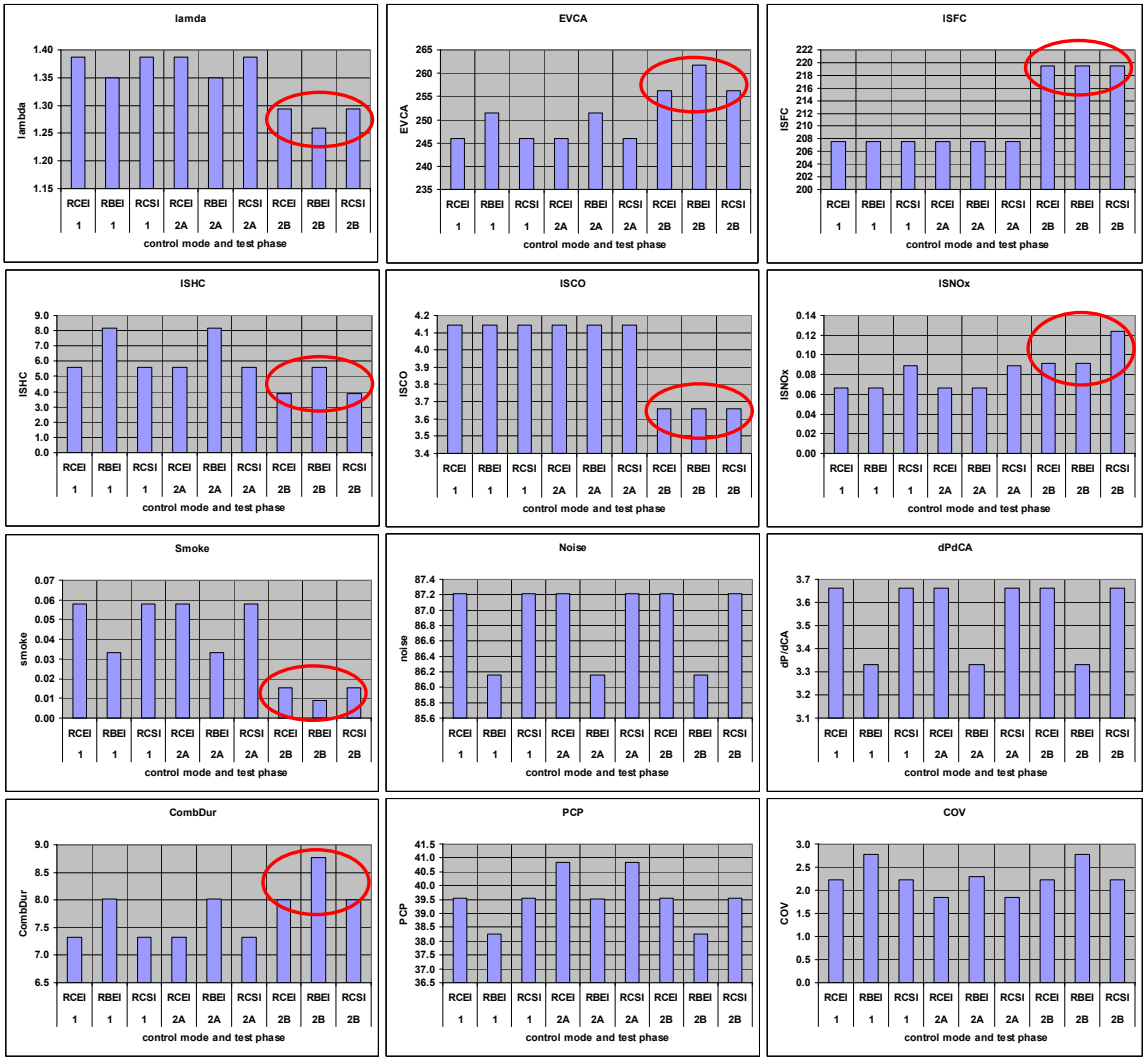


Figure 8.3: Test Phase and Control Mode Effects for Average Fuel at 2000 rpm Condition.

8.4 Fuel Variables That Affect Engine Performance

The main goal of the AVFL-13 and 13b projects was to study the effects of fuel chemistry and properties on HCCI combustion. A broad correlation study was conducted over all the speeds and control modes and using the 249 parametric fuels described above. Table 8.4 indicated which engine responses show correlation to fuel variables, with yes indicating an r value ≥ 0.7.

Table 8.4: Engine Performance Variables that are Affected by Fuel Characteristics

	1000 rpm	2000 rpm	3000 rpm
lambda	yes	yes	yes
EVCA	yes	yes	yes
ISFC	no	yes	yes
ITE	no	yes	yes
ISHC	no	no	yes
ISCO	yes	yes	no
NOx	no	yes	yes
Smoke	yes	yes	yes
Noise	no	yes	yes
dPdCA	yes	yes	yes
CombDur	no	yes	no
PCP	yes	yes	no
COV	yes	yes	yes
CombEff	no	yes	no
ExhT	no	no	no

The number of combinations for which correlation graphs can be made is enormous, since there are 16 fuel variables (seven in the models and seven more predicted from the nine), 15 engine response variables, three operating modes (speeds), three control modes, and effects of ethanol splash blending. If one also considers four levels for ethanol blending (0, 10, 20, and 30%), this multiplies out to a total of 8640 engine response / fuel parameter correlations. In this report, the task has been simplified by focusing on only a subset of these relationships. The following discussion focuses mainly on RCEI control mode (previously identified as best overall mode), looks for similar responses in two of three operating modes (speed), and focuses mainly on EVCA, ISFC, NOx, Smoke, and dP/dCA as the main variables of interest and as variables which can be used to indicate response of other variables. This study also found that vector value V1 correlates mainly with ISFC and PCP, V2 correlates mainly with EVCA, NOx, dP/dCA, and COV, and Smoke correlates mainly with V3. These relationships can be further decoded to individual fuel chemistry or properties using the equations embedded in the fuel panel simulator.

As an example of this decoding, it was found that ISFC correlated to vector value 1; the results are shown in Figure 8.4. This graph indicates an increase in ISFC as vector 1 value increases for 2000 and 3000 rpm, but not for 1000 rpm. The graph also shows that 2000 rpm is a more efficient speed than 3000 rpm.

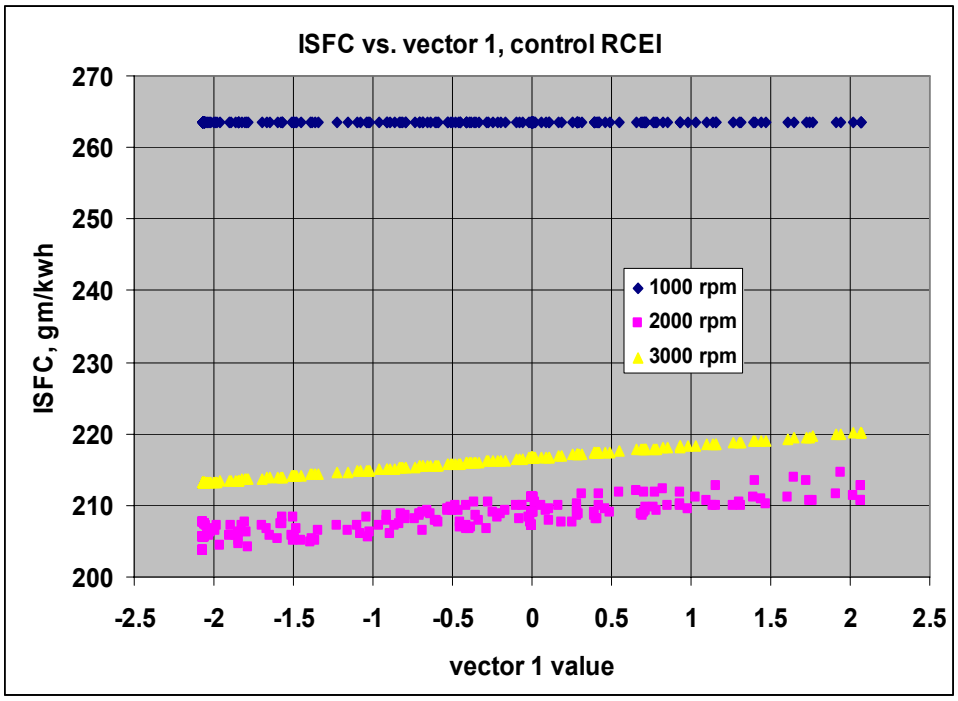


Figure 8.4: Relationship of ISFC to Score for Fuel Vector 1.

This change in fuel economy with vector 1 may be due to an energy content change, to a change in how the engine runs or is controlled, or to other fuel related factors. There is not a simple relationship between ITE and ISFC (see Figure 8.5), indicating that the effect is more than can be explained by energy content. Further investigation indicates that ISFC is hurt by things that decrease gravimetric basis energy density (i.e., BTU/lb), such as higher aromatics, high boiling points, and/or high specific gravity, and that ISFC is also worse with high MON, low sensitivity fuels. These fuels are harder to ignite and require an earlier exhaust valve closing angle and a correspondingly lower lambda.

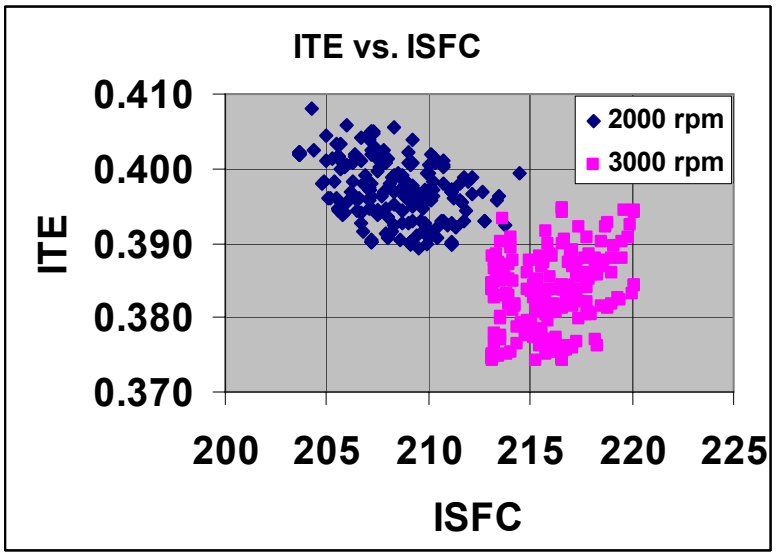


Figure 8.5: Comparison of ITE and ISFC for Fuel Variations, Indicating that Changes are Not Simply Due to Energy Content Changes.

Fuel vector 2 shows a strong influence on ignition, as indicated by the relationship between EVCA and V2. A higher value or later EVCA indicates easier ignition and this corresponds to higher values for V2. This relationship is shown in Figure 8.6.

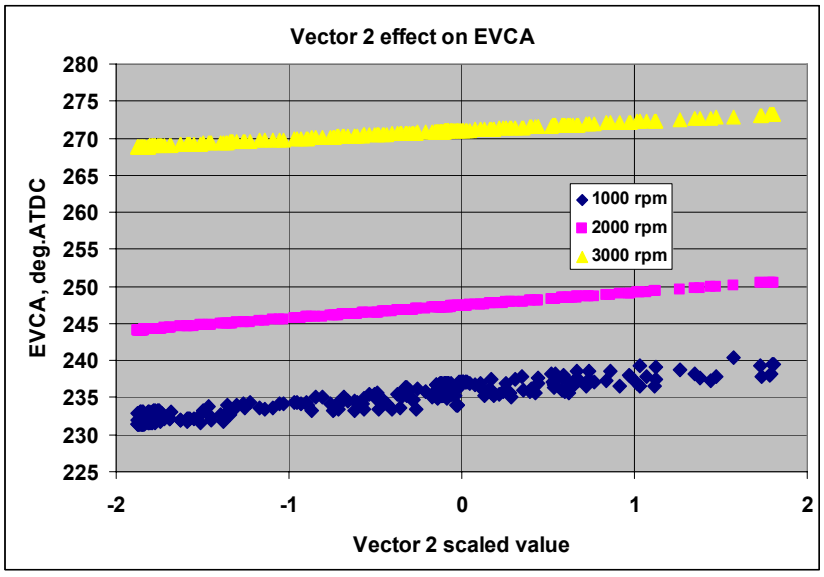


Figure 8.6: Dependence of EVCA on Fuel Vector 2 Score

EVCA actually correlates best to MON and sensitivity, as shown in Figure 8.7, both of which would be expected to influence ease of ignition, with lower MON, higher sensitivity fuels

showing later EVCA. MON always shows better correlation to EVCA than RON, and sensitivity shows stronger correlation than MON at 1000 RPM. This type of engine is generally considered to be MON driven or 'beyond MON', meaning that combustion is primarily driven by heat rather than pressure. 'Beyond MON' means that the effective octane of the fuel at ignition is lower than the MON rating of the fuel due to the higher temperatures at ignition compared to the octane rating engine. The 1000 rpm condition is farther 'beyond MON', because of earlier EVCA and heated intake, and therefore sensitivity plays a larger role.

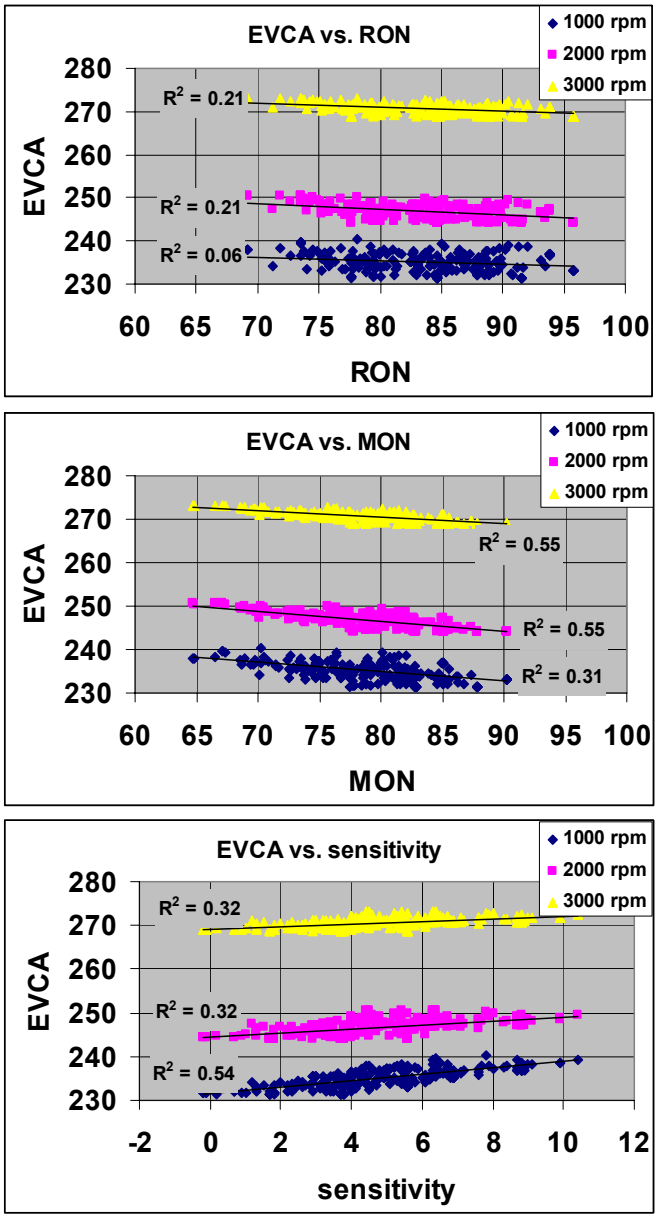


Figure 8.7: Dependence of EVCA on RON, MON, and Sensitivity for Parametric Fuel Set.

NOx correlates well to vector 2 value for 2000 and 3000 rpm, as shown in Figure 8.8. There is no correlation indicated for 1000 rpm, but NO_x for that operating condition is 0.01 gm/kg fuel or

0.1 ppm, so it is essentially at the noise level of the instrument. NO_x also shows some correlation to vector 3, which is a reason for the width of these relationships.

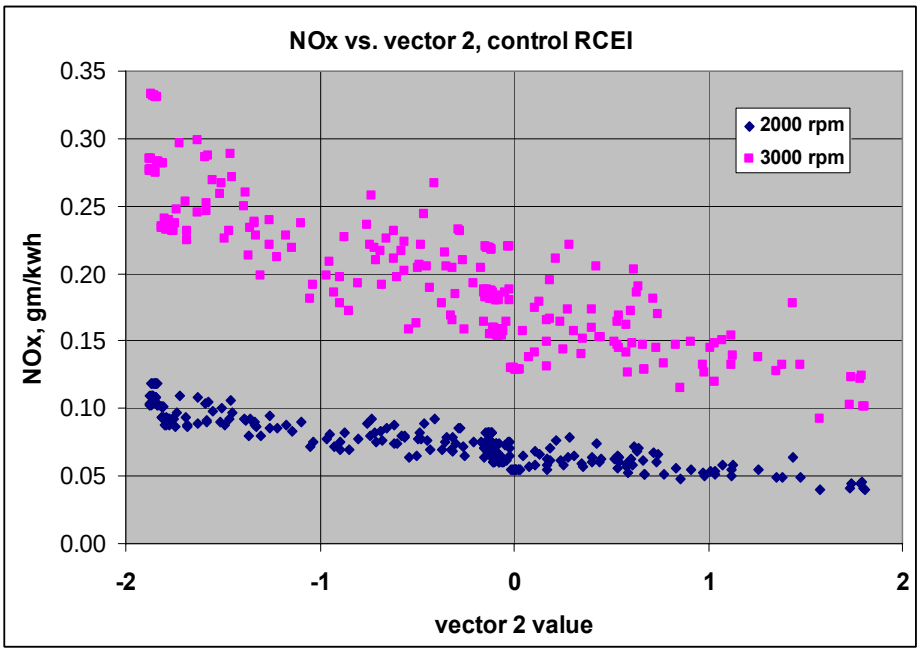


Figure 8.8: NO_x Dependence on Vector 2 Value for Parametric Fuel Set.

NOx is often found to be correlated to peak combustion temperatures, which are not available in the experimental data. However, peak temperature can also correlate to dP/dCA, the maximum rate of cylinder pressure rise. Figure 8.9 indicates that dP/dCA also correlates with vector 2 value for 2000 rpm. There is no correlation at 1000 rpm, perhaps because peak pressure occurs after TDC and is lower than compression pressure. There is no correlation for 3000 rpm because dP/dCA was controlled to a constant value of 5.5 bar/degree.

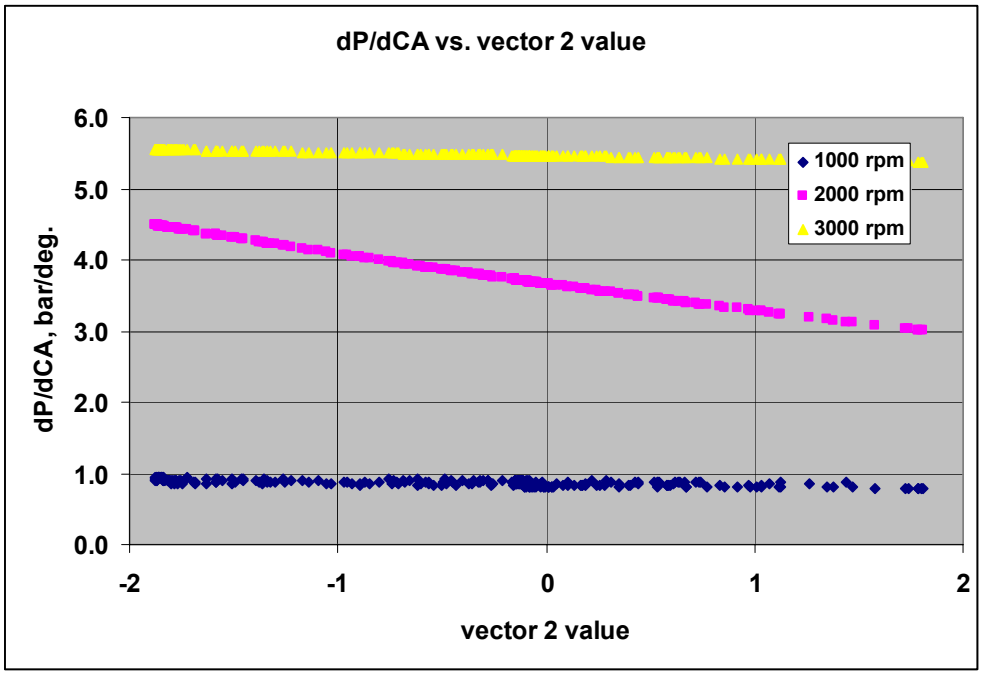


Figure 8.9: dP/dCA Dependence on Vector 2 Value for Parametric Fuel Set.

NOx also correlates to fuel MON. Higher MON fuels produce more NO_x because they are harder to ignite, require early EVCA, and this results in higher dP/dCA. These relationships are shown in Figure 8.10.

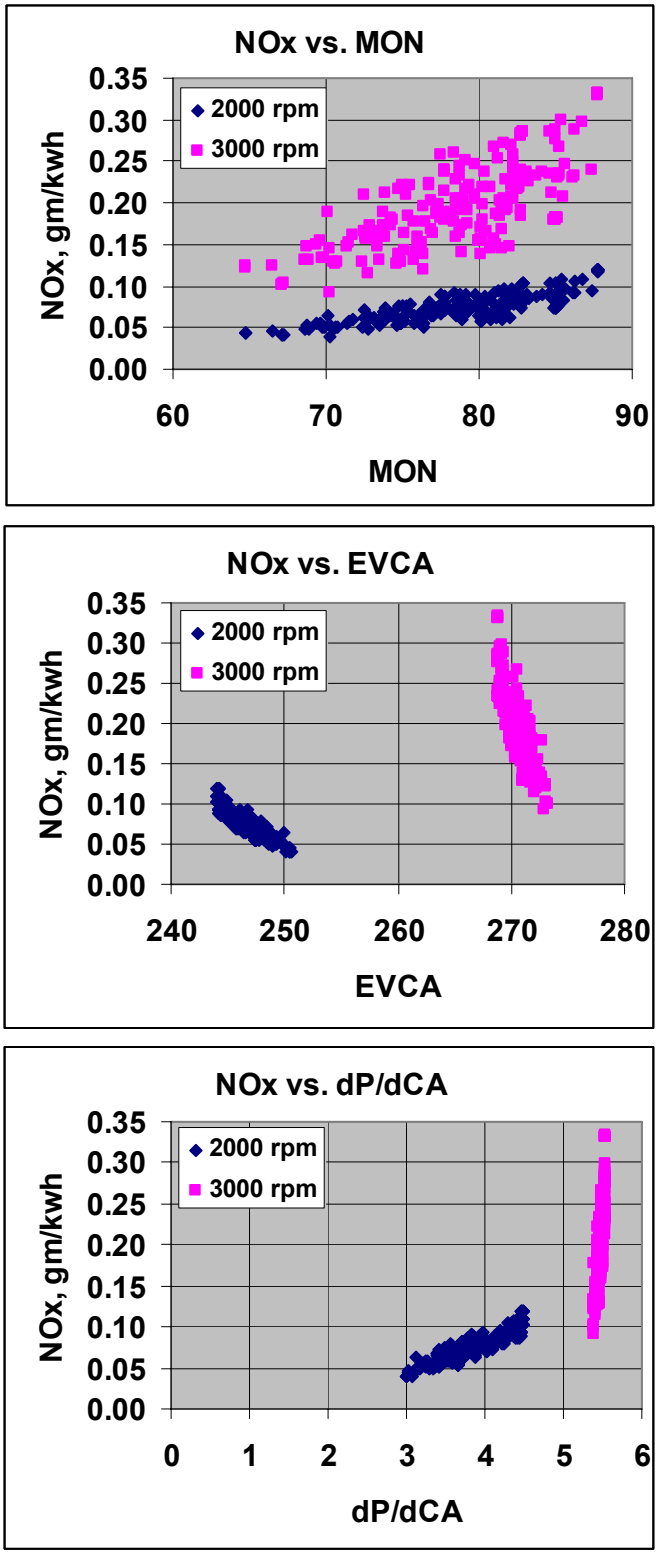


Figure 8.10: NO_x Dependence on MON, EVCA, and dP/dCA.

Smoke relates most strongly to vector 3, with some trends to vector 1. The relationship between smoke and vector 3 is shown in Figure 8.11, for each operating mode. In all cases, smoke increases as vector 3 decreases, with the trends being stronger for 2000 and 3000 rpm.

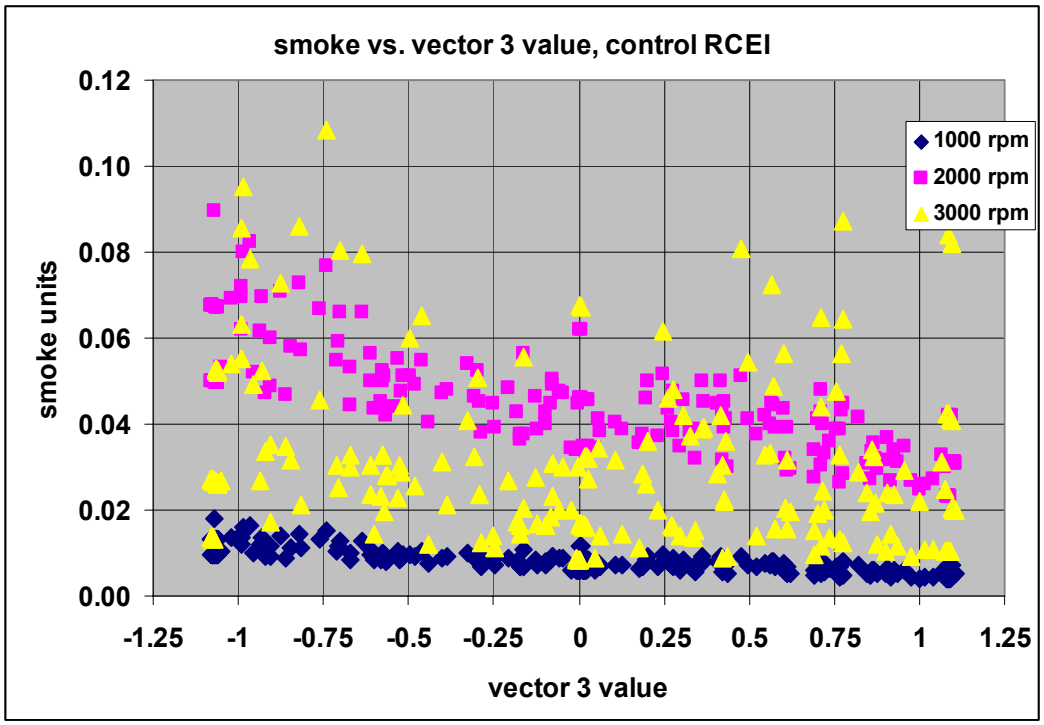


Figure 8.11: Smoke Dependence on Vector 3 Value for Parametric Fuel Set.

Smoke can also be compared to actual fuel properties and chemistry. Smoke is found to relate to aromatic content, specific gravity, and distillation temperatures and an increase in any of these is found to increase smoke. 3000 rpm is most sensitive for smoke and 1000 rpm is least sensitive. Smoke values for all conditions are low, but the trends towards higher smoke for heavier, higher boiling, more aromatic fuels are consistent. These relationships are shown in Figure 8.12.

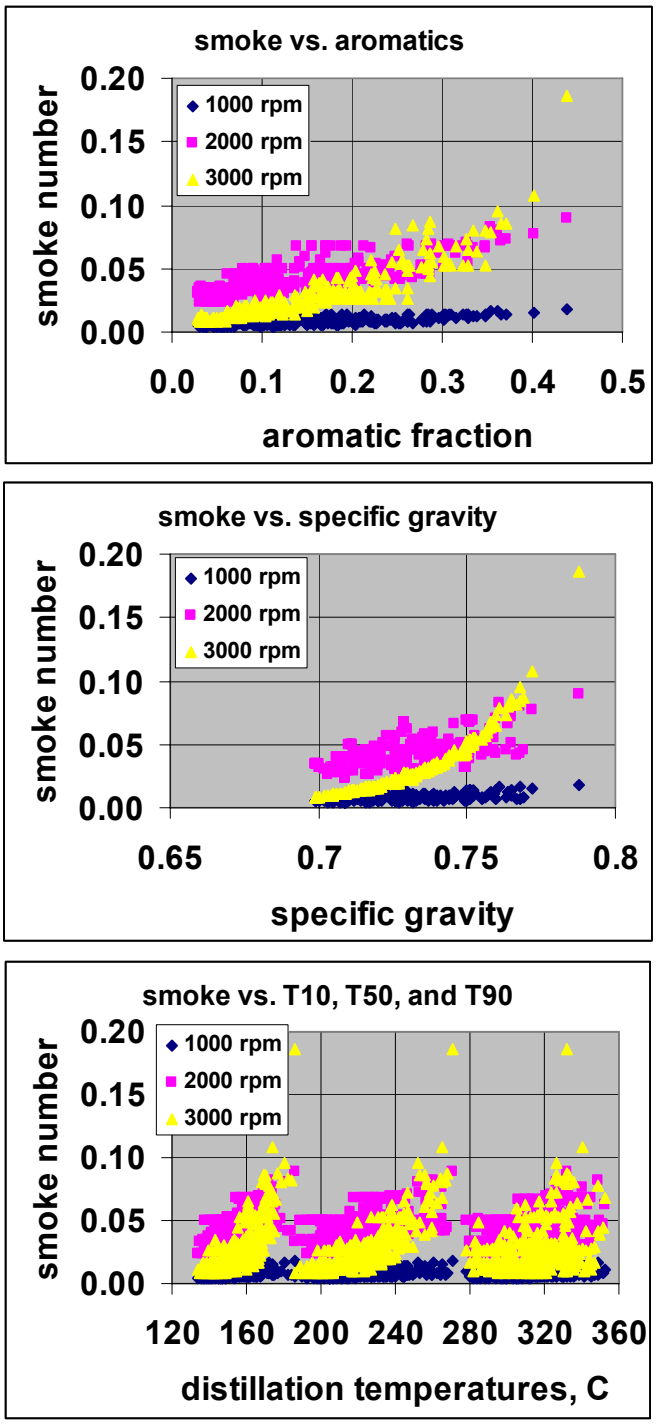


Figure 8.12: Smoke Dependence on % Aromatics, Specific Gravity, and Distillation Temperatures.

8.5 Desirable Fuel Variables for Overall Engine Performance

If criteria are devised to sort engine data by fuel characteristic, then fuels can be identified that provide improved performance for this engine and the selected operating and control conditions. Inspection of the previous figures suggests that fuels with a low value for vector 1 and high values for vector 2 and 3 would provide improved ISFC, lower smoke and NOx, lower dP/dCA, and easier ignition. However, this 'optimization' applies only to this engine, operating conditions, and fuel set and should not be considered a universal best choice for advanced combustion engines and all operating conditions. The 249 fuel parametric fuel set described previously was sorted by vector values and fuels were selected at the desired ranges of vector values. The desired vector value ranges were set to provide some fuels which met all three criteria. This process yielded three fuels which averaged: 12% n-paraffins (excluding C_4); 2% n-C_4; 35% iso-paraffins; 15% cyclo-paraffins; 25% olefins; 11% aromatics; 152°C T10; 219°C T50; 294°C T90; 78 RON; 71 MON; and 7 sensitivity.

This describes a fuel with low octane, high sensitivity, low aromatics, and in the lower range of distillation temperatures and density. These fuels would ignite more easily, with a higher lambda and later (higher) EVCA and produce lower smoke. They are best represented by a blend of C100 (olefins) for high sensitivity with low aromatics and D100 (straight run) for low octane. The desired ranges of vector values are shown in Figure 8.13, as the boxes formed by the intersection of the ranges of vector values desired. Also shown on this plot are fuel C100, C85D15, and C50D50. It appears that a fuel blend between C85D15 and C50D50 falls into the desired range.

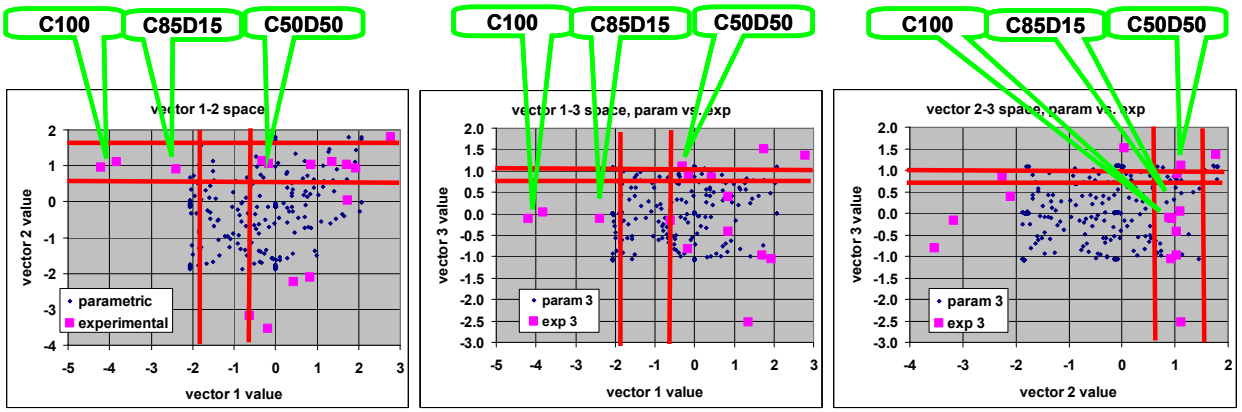


Figure 8.13: Vector Ranges and Fuels which Provide Desirable Operating Characteristics.

8.6 Desirable Fuels for Constant (R+M)/2

The selection of desirable fuels for HCCI engine operation above assumes that it would be possible to change octane values significantly, which would essentially produce a new gasoline fuel of low octane (78 RON and 71 MON). This section examines the possibility of selecting chemistry or properties to produce improved HCCI operation if choice is restricted to octanes meeting today's criteria. This parametric study uses fuels of (R+M)/2 equal to 87, which is the

current requirement for unleaded regular. This octane was selected because it is closer to the midpoint of the fuel set octane and will provide more options for varying chemistry. This octane can be met with a tradeoff of the chemistries of aromatic, olefin, and iso-paraffin to build octane and n-paraffin to trim octane value. This fuel set was described above, as the last fuel set in section 8.1. Olefins and aromatics can be adjusted to provide high sensitivity fuels and iso-paraffins can be used to provide low sensitivity fuels. N-paraffins are relatively constant across all the fuels (see Table 8.2) because of the constant octane target. The 32 fuels in Table 8.2 cover sensitivities of 0 to 11, olefins from 0 to 29%, aromatics from 3 to 44%, and iso-paraffins from 28 to 85%. These fuels are used as inputs to the engine simulator models, which are otherwise set to nominal test mode conditions, 0% ethanol, and RCEI control mode. For these fuels, the total combined amount of olefins and aromatics determine sensitivity level, but the ratio between olefins and aromatics can be varied at a given sensitivity level. This relationship is shown in Figure 8.14, which shows the olefin and aromatic fraction for the three sensitivities in the parametric fuel set. The fuels with essentially no sensitivity have all iso-paraffins and almost no olefins and aromatics. The mid-sensitivity fuels allow the most option for trading between olefins and aromatics, and the high-sensitivity fuels begin to run out of options because of chemistry limits in the blending components.

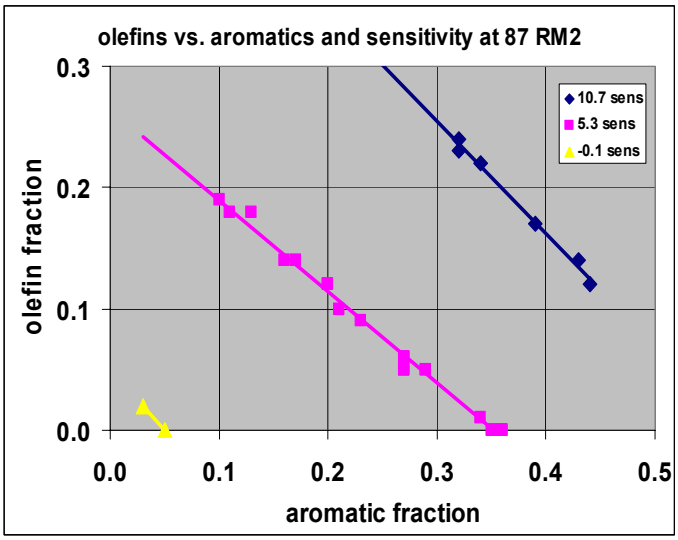


Figure 8.14: Tradeoffs of Olefins and Aromatics for 87 Octane (R+M)/2 Fuels of Different Sensitivity Levels.

The 31 fuels were used as inputs to the three engine simulators, which were otherwise set for RCEI control and a mid-point test phase. For each speed, correlation tables were prepared for engine response and fuel parameters, and Table 8.5 is one such table, shown for 2000 rpm. There are a large number of correlations, highlighted when $r \geq 0.7$. All three correlation tables were examined, and visualization of selected engine responses will be made, selecting those which were highlighted previously for fuels covering the entire range of fuels in the experiments.

Table 8.5: Correlations of Engine Response to Fuel Parameters for (R+M)/2 = 87 at 2000 rpm.

	lamda	EVCA	ISFC	ITE	ISHC	ISCO	NOx	Smoke	Noise	dPdCA	CombDur	PCP	COV
ParaXC4	0.53	0.84	0.64	0.55	0.00	0.84	0.69	0.74	0.84	0.84	-0.84	-0.63	0.84
ParaC4	-0.97	-0.86	0.31	-0.98	0.00	-0.86	0.19	0.13	-0.86	-0.86	0.86	-0.32	-0.86
iPara	-0.87	-1.00	-0.21	-0.88	0.00	-1.00	-0.31	-0.38	-1.00	-1.00	1.00	0.20	-1.00
cycloPara	0.94	0.71	-0.55	0.92	0.00	0.72	-0.48	-0.41	0.71	0.71	-0.71	0.56	0.72
Olefins	0.87	0.56	-0.71	0.85	0.00	0.57	-0.64	-0.57	0.56	0.56	-0.56	0.72	0.57
Aromatics	0.47	0.82	0.69	0.51	0.00	0.81	0.76	0.80	0.82	0.82	-0.82	-0.69	0.81
T10	-0.33	0.12	0.99	-0.30	0.00	0.11	0.97	0.96	0.12	0.12	-0.12	-0.99	0.11
T50	-0.14	0.32	0.98	-0.11	0.00	0.31	0.98	0.98	0.32	0.31	-0.32	-0.98	0.31
T90	-0.92	-0.65	0.65	-0.90	0.00	-0.65	0.56	0.50	-0.65	-0.65	0.64	-0.66	-0.65
V1	-0.55	-0.13	0.95	-0.53	0.00	-0.13	0.91	0.88	-0.13	-0.13	0.13	-0.96	-0.13
V2	0.89	1.00	0.16	0.90	0.00	1.00	0.26	0.33	1.00	1.00	-1.00	-0.15	1.00
V3	-0.40	-0.77	-0.75	-0.43	0.00	-0.76	-0.80	-0.84	-0.77	-0.76	0.77	0.74	-0.76
RON	0.92	0.99	0.07	0.93	0.00	0.99	0.17	0.25	0.99	0.99	-0.99	-0.06	0.99
MON	-0.93	-0.97	-0.02	-0.94	0.00	-0.98	-0.11	-0.19	-0.97	-0.97	0.97	0.01	-0.98
sens	0.94	0.99	0.04	0.95	0.00	0.99	0.14	0.21	0.99	0.99	-0.99	-0.03	0.99
RM2	0.03	0.13	0.19	0.06	0.00	0.12	0.24	0.24	0.13	0.13	-0.14	-0.19	0.12

The first relationship, shown in Figure 8.15, is between EVCA and sensitivity and indicates that higher sensitivity fuels ignite more easily, even at constant (R+M)/2. Fuels which ignite more easily have a later (larger) value for EVCA and also will have a higher lambda value. Engine conditions with higher lambda generally have improved ISFC, but the effects under the conditions of this particular study, at constant (R+M)/2, are very small.

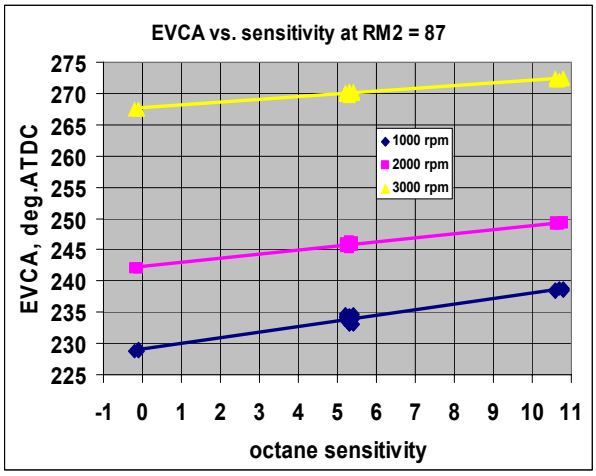


Figure 8.15: EVCA vs. Sensitivity for (R+M)/2 = 87.

Figure 8.14 showed the linear relationship between olefins and aromatics in these fuels and the ability to trade olefins and aromatics at constant sensitivity. The Figures below, Figure 8.16, indicate that higher aromatic fuels exhibit both higher NO_x and smoke than those with lower aromatics (i.e., higher olefins). The actual NO_x and smoke values are also small, but trends might be useful in helping to chose fuel blend formulas.

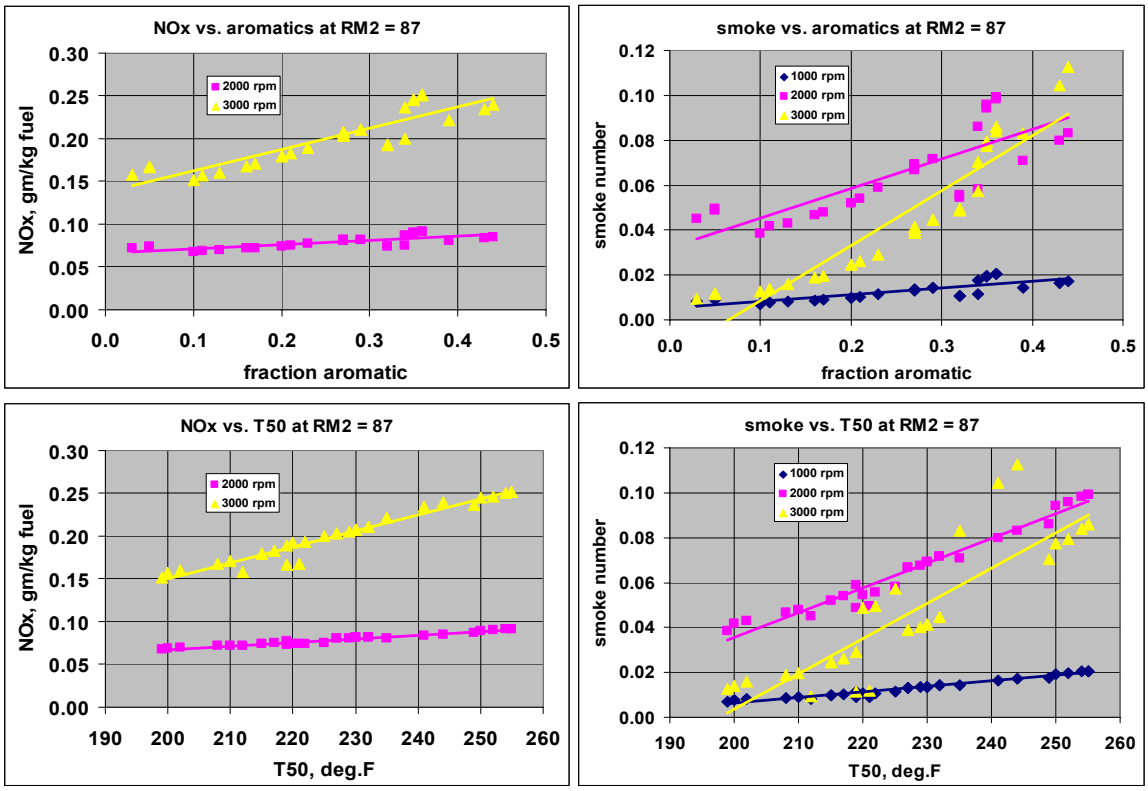


Figure 8.16: Relationship of Smoke and NO_x to Aromatics for (R+M)/2 = 87.

Overall, fuels with high sensitivity, low aromatics, and low T50 improved engine performance in terms of ISFC, NOx, and smoke, even when (R+M)/2 is held constant.

8.7 Effects of Ethanol Blending

Ethanol was splash blended into selected base fuels without consideration of RVP or other fuel property controls. As described in the fuel model section of this report (section 4), statistical modeling of the data was difficult because of the limited number of fuels, because the ethanol fuels were run in two phases (2a and 2b) with a major engine rebuild between, and because the addition of the ethanol fuels resulted in a major departure from the initial experimental plan. In the study described in this section, ethanol content was blended to the 249 fuel parametric fuel set, by adding an input line in the engine simulators varying from 0 to 30% ethanol blend level, on a random basis. The simulators were otherwise set for RCEI operation and an average test phase. This manner of constructing the fuel inputs to the simulator would be expected to answer the question, "over the entire range of fuel properties and chemistry of the base fuels, what additional engine operation effects are predicted due to ethanol blending". This study could also be done with a specified base fuel chemistry, but it is felt that the model is not robust enough to provide answers to this level of detail. This is because of the entanglement of ethanol results with the engine rebuild between test phases 2a and 2b. The engine simulator outputs were collected and results correlated to % ethanol content. Some engine response variables showed good correlations at some operating conditions, and others did not. These correlations are shown in Table 8.6. No trends extend across all three operating conditions (speeds) and few extend across even two conditions.

Table 8.6: Correlation of Engine Response Variables to % Ethanol in Fuel Blends.

	1000 rpm	2000 rpm	3000 rpm
lambda	0.06	0.62	0.06
EVCA	0.06	0.03	0.03
ISFC	0.00	-0.04	0.94
ITE	0.00	0.08	0.03
ISHC	0.00	1.00	0.01
ISCO	0.47	0.03	0.00
NOx		-0.03	-0.57
Smoke	0.01	0.01	0.54
Noise	-0.09	-0.03	-0.32
IMEP			0.00
dPdCA	0.03	-0.04	-0.03
CombDur	1.00	0.03	0.00
PCP	0.08	-0.91	-0.97
COV	0.09	0.03	0.02
CombEff	-1.00	-0.08	0.00
ExhT	-1.00	0.00	-1.00

Interpreting these results is difficult because of the lack of overall trends. The discussions which follow include both an interpretation of Table 8.6 and engineering judgment and logic applied to expected results. The interpretations are offered in attempt to gain something general from the data and to help guide future studies. First, ethanol blending appears to hurt ISFC, but not ITE. This appears to be an energy content issue, as shown by the parallel trends of loss in ISFC and loss in energy content with ethanol blending, shown in Figure 8.17 for 3000 rpm. In this figure, the ISFC response also shows a non-linear effect, but since this is introduced by results from only one data point, it is hard to put too much confidence in this trend. Figure 8.17 does show that the overall losses in ISFC and energy content and are of similar magnitude.

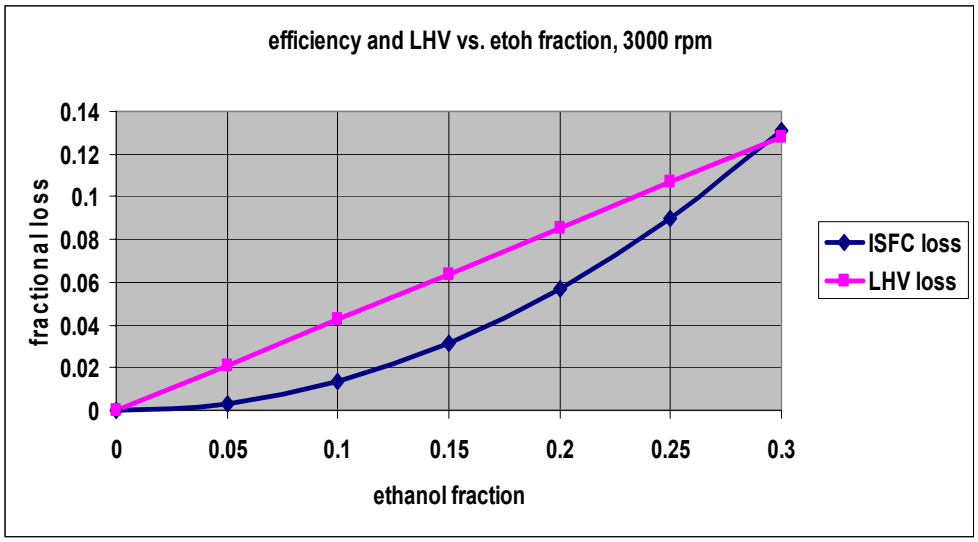


Figure 8.17: Fractional Loss of ISFC and Energy Content with Ethanol Blending at 3000 RPM.

Other observations which can be made are as follows. Ethanol blending does not appear to affect HCCI ignition as indicated by EVCA and lambda, perhaps because the increasing octane effect is counteracted by an increased fuel sensitivity and oxygen content. Ethanol blending appears to increase HC, CO, and/or smoke while lowering NOx, exhaust temperature, noise and peak cylinder temperature. This may be due to cooler cylinder conditions resulting from ethanol evaporation, but data are not available to support this observation. Overall, the trends from ethanol blending appear logical and it appears that ethanol blended fuels are acceptable for this type of HCCI engine up to the 30% blend levels tested.

9. DESCRIPTION OF FUEL AND ENGINE SIMULATORS

The fuel and engine simulators developed in this work have been supplied in the form of Excel spreadsheets (2003 version) that can be exercised by the user to perform fuel and engine performance studies like those described in prior sections. Because the simulators contain all of the predictive models developed in the work, including required data and mathematical formulas, they also serve as the primary documentation of those models. The following sections describe the simulators and give instructions for their use. The user is cautioned to save models to a different name before inputting data, since it is fairly easy to disrupt the active cells by edit commands until one is completely familiar with the models.

The user is responsible for selection of input data and for their appropriateness for use with the simulators. The structure of the simulators is that they will mechanically translate input as specified into outputs, with minimal range checks, limitations or warnings to the user. In the studies conducted for this report, considerable attention was given to keeping the inputs within the range of the experiments that were conducted and to discount predictions of performance metrics that approached or exceed their experimental limits. If the user inputs implausible or out-of-range data, the simulators should be expected to produce implausible or out-of-range predictions for performance.

9.1 Fuel Simulator

Excel spreadsheet FuelPanel v20091231.xls implements the complete fuels model developed in this work. It was termed the FuelPanel because the first tab provides an interactive tool that allows the user to create fuels based on assumed vector characteristics, much like adjusting knobs on a control panel, and by doing so, to begin to understand how the vectors are related to the chemical and physical properties of fuels. The second and third tabs of the spreadsheet provide means for operating the fuel model in both forward and reverse directions – i.e., to specify a fuel in terms of chemistry, boiling points, and ethanol content and have its vectors calculated (the forward direction), or to specify a fuel in vector terms and have its chemistry, boiling points, and other properties computed (the reverse direction).

9.1.1 Fuel Panel Tab

The Fuel Panel tab is intended as an educational tool. Five slider controls are present that may be used to set the Vector 1-4 scores to define a base fuel and the ethanol content that is optionally blended into the base fuel. These five values fully define a finished fuel for the purposes of the analysis of fuel properties in the fuel model. When the sliders are clicked once on the right or left arrows, the vector score will increment by 0.1 units and the ethanol content will increment by 1%. A larger move is made by clicking the open space between the central bar and the right and left arrows; the vector score will increment by 0.5 units and the ethanol content will increment by 5%. The user may also click on and hold the central bar and move it to any point within the range. By using the slider controls in these ways, the vector scores and fuel ethanol content can be varied between the minimum and maximum values encountered in the experimental fuels.

As the input values are changed, the graphs of fuel properties will change automatically to correspond, although the rate at which the graphs update may depend on the speed of the computer. Figure 9.1 shows an example of the graphical output. This particular example shows the properties of the average base hydrocarbon fuel with no ethanol content. This fuel would be specified as scores of zero for Vectors 1-4 and as having 0% ethanol content. Ten of the properties displayed are used in the specification of the base fuel vectors (the six values for chemistry and the three boiling points) or of the splash blending done to create an oxygenated fuel (ethanol content). The displayed properties are predictions made from the specification of a fuel using the procedures described in Section 4 – RON, MON, sensitivity (RMSens), specific gravity, and RVP. The RVP of base hydrocarbon fuels cannot be predicted with information contained in the dataset and is set equal to the average 7.1 psi value (unless modified for oxygenated fuels by the added ethanol content).

Cautions: only the five sliders should be used to set the vector score and ethanol content values. Typing a value directly into the box cells holding those values will disconnect the slider from the formulas.

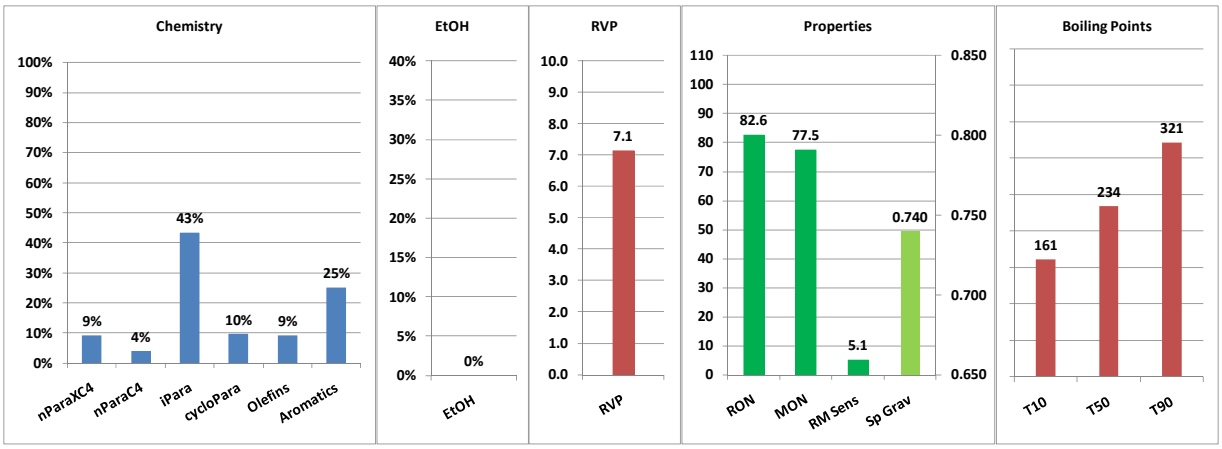


Figure 9.1: Graphical Output of the Fuel Panel (example for average fuel)

9.1.2 fromFuelSpec Tab

The fromFuelSpec tab operates the PCA process in the forward direction, taking the input specification of a fuel in terms of six chemistry variables, three boiling points, and the presence (if any) of ethanol and producing, as outputs, the full slate of fuel chemistry and properties for the finished fuel and the scores for Vectors 1 through 4. The input area is structured to allow up to 100 fuels to be input and evaluated. The spreadsheet comes with the 24 experimental fuels used in the analysis.

From the specification of the base fuel and ethanol content (if any), the spreadsheet computes these characteristics of the finished fuel:

- The six chemistry variables and ethanol content. For oxygenated fuels, the chemistry values have been diluted volumetrically to account for the splash blending of the specified volume of ethanol. For non-oxygenated fuels, the chemistry values are those input by the user to specify the base fuel.
- Eight physical properties: T10, T50, T90, RON, MON, RMSen, SpGrav and RVP. For non-oxygenated fuels, the boiling points are the values input by the user to specify the base fuel.
- The scores for Vector 1, 2, 3, and 4. These are the vectors that are used in the predictive models for base fuel properties and as independent fuel variables in the predictive models of engine performance. However, they are only four of the 9 vectors in the full vector expression for fuels; all nine vectors would be required to exactly replicate the input fuel.

The fuel input area is structured in the way that fuels are input to the engine simulators. If fuel specifications are available from another source, perhaps a fuels survey or other experiments, the data that is input to this tab to evaluate fuel properties can also be transferred directly to the engine simulators using a copy/paste operation.

Cautions: only the cells displaying in red font (B5:J104) should be changed by the user. Data may be entered in this area by typing values or by using a copy/paste special/values operation. The paste special/values choice is recommended so that only the data content of cells is

modified, leaving cell formatting unchanged. The user should avoid inserting or deleting rows or sorting the data area, because doing so has the potential to disrupt active formulas in the sheet. It is best to perform data manipulations in an inactive sheet elsewhere.

9.1.3 fromVectors Tab

The fromVectors tab operates the PCA process in the reverse direction, beginning with the input specification of a fuel in terms of its base fuel scores for Vectors 1-4 and the volume of ethanol (if any) that is blended. From these five variables, the spreadsheet estimates the six chemistry variables, three boiling points, and the five additional properties RON, MON, RMSens, SpGrav and RVP. The input area is structured to allow up to 100 fuels to be evaluated. The spreadsheet comes loaded with the 5-variable representation of the experimental fuels.

The term “estimate” was used for a purpose in stating that this tab estimates chemical and physical properties of fuels. Vector 1 through 4 of the fuel model are an adequate representation of fuels for the analytical purposes of this study. However, all 9 vectors defined by PCA would need to be carried in the calculations if one were to exactly replicate the fuels. Thus, a representation in terms of four vectors and ethanol content is an approximation of the fuel, although a very close one. One can see the degree of difference by comparing the output of this tab to the input fuel specifications in the fromFuelSpec tab, from which the vectors where taken. In most cases, chemistries differ by only a few percentage points and boiling points by only a few degrees Fahrenheit.

This tab can be used to evaluate any matrix of vector and ethanol choices created by any means, but it is provided mainly to support parametric fuel studies. Parametric studies are ones in which fuels are identified by sampling the vector fuels space either systematically (using discrete levels) or through a random balance process. By sampling according to vectors, one maintains the correlations among individual fuel properties that were present in the experimental fuels. The methods of parametric study were discussed in Section 4. To perform a study using systematic sampling, one would create the nested itemization of all possible levels in another location and transfer the resulting matrix of vector and ethanol choices into the data entry area.

The user faces two analytical decisions in using this tab. The first decision is whether, in specifying vector values and ethanol content, the choices should be constrained to lie within the range of the experimental fuels, or not. Staying within the experimental range maintains the greatest fidelity with the experiment, but there may be circumstances where it is useful to explore outside the range. The second decision is whether the resulting parametric fuels should be further screened before use in an engine simulator. While the vector fuel space is based on the characteristics of the experimental fuels, it is not the case that the experimental fuels necessarily explore all of that space. Fuels created by vector sampling are ones that could be created through the blending process used to create the experimental fuels. There is no guarantee, however, that all of the fuels would meet gasoline-range fuel specifications or fall within the specific range explored by the experimental fuels. Again, staying within the experimental range maintains the greatest fidelity with the experiment, but there may be circumstances where it is useful to explore outside the range. The user is responsible for the consequences of these choices. The data given in Table 9.1 on the experimental range of the variables involved may be helpful in formulating parametric studies.

Cautions: only the cells displaying in red font (B5:F104) should be changed by the user. Data may be entered in this area by typing values or by using a copy/paste special/values operation. The paste special/values choice is recommended so that only the data content of cells is modified, leaving cell formatting unchanged. The user should avoid inserting or deleting rows or sorting the data area, because doing so has the potential to disrupt active formulas in the sheet. It is best to perform data manipulations in an inactive sheet elsewhere.

9.1.4 Other Spreadsheet Tabs

Two other tabs are present in the fuel simulator. Tab PCA summarizes the parameters of the PCA analysis of base fuels. These values are used in computing the vector representation of fuels in both the forward and backward directions. Tab Predictive Models summarizes the coefficients and mathematical forms of the two-stage predictive model for other fuel properties, which was described in Section 4.3.

9.2 Engine Simulators

The three Excel™ spreadsheets:

- AVFL13(AB)_Simulator_1000 v2009.12.31.xls
- AVFL13(AB)_Simulator_2000 v2009.12.31.xls
- AVFL13(AB)_Simulator_3000 v2009.12.31.xls

Table 9.1: Experimental Fuel Range Covered by AVFL-13 Program.

	Minimum Value	Maximum Value
Vector and EtOH Inputs		
Vector 1	-4.2	2.8
Vector 2	-3.5	1.8
Vector 3	-2.5	1.5
Vector 4	-1.1	0.8
EtOH	0%	35%
Fuel Property Values		
ParaXC4	0.00423	0.1962
ParaC4	0.0004	0.0670
iPara	0.2722	0.8865
cycloPara	0.0110	0.1693
Olefins	0.0033	0.4470
Aromatics	0.0280	0.5495
T10	133.7	191.1
T50	174.6	271.8
T90	239.7	354.0
EtOH	0%	35%

are the engine simulators for operating Modes 1 (1000 rpm), 2 (2000 rpm), and 3 (3000 rpm), respectively. The simulators come with the complete experimental datasets in two series, one for base fuels and another for oxygenated fuels.

Using the Inputs tab, the user specifies the conditions of up to 150 engine tests by the values given to these variables in the input area (red font) in Columns S through AI:

- EV2 = 1 to indicate the presence of a secondary exhaust value opening (RBEI control mode), or the value zero otherwise (other modes)
- Inj2 = 1 to indicate the presence of a secondary injection invent (RCSI control mode), or the value zero otherwise (other modes)
- Nine values to specify the fuel, using the six chemistry values and three boiling points. These inputs are defined for the base hydrocarbon fuel involved, even when the test pertains to an oxygenated fuel.
- The ethanol content value for testing use an oxygenated fuel.

- Three dummy variables representing departures from the experimental targets for engine output, combustion timing, and intake air temperature. As provided, these columns contain non-zero values to represent that actual test conditions in the experiment. Normally, these values will be set to zero in studies in order to remove the effects of experimental imprecision from the predictions of engine performance.
- Two dummy variables representing the program Phase in which the test was conducted. As provided these columns contain one or zero values to represent the actual experimental data. The variable Phase2 takes on the value 0 for all Phase 1 tests, and the value 1 for all Phase 2 tests. The variable Phase 2b takes on the value 0 for all Phase2a tests, and the value 1 for all Phase 2b tests. . Except where engine performance in a specific phase is of interest, these variables should be set to the values Phase2=0.31 and Phase2b=0.15 to control the predictions to represent an average across phases for the program.

The Inputs tab indicates the units to be used in specifying the inputs and also gives the range of values for these variables encountered in the experimental data.

Note that the input area specified fuels in terms of the familiar chemical and physical variables that make up the fuel vectors, rather than the vector scores themselves. This structure was adopted so that fuels data from any appropriate source could be used in the simulators without the need for the intervening step of computing fuel vectors. However, the models have only been verified for fuels covering the range and characteristics of the experimental fuels.

The model predictions for engine performance are presented in Columns B through Q of the Inputs tab. Sixteen different performance metrics are predicted:

- Lambda and EVCA – the air fuel ratio (lambda form) and exhaust value closing angle required to achieve the experimental targets of engine output and combustion phasing
- ISFC and ITE – indicated specific fuel consumption and indicated thermal efficiency
- ISHC, ISCO, and NO_x – emissions of hydrocarbons and carbon monoxide (gm/kg Fuel) and NO_x (ppm)
- Smoke and Noise
- IMEP and dPdCA – indicated mean effective pressure (bar) and peak rise in cylinder pressure (bar/degree)
- CombDur – combustion duration
- PCP – peak cylinder pressure
- COV – coefficient of variation for IMEP
- CombEff – combustion duration
- ExhT – exhaust gas temperature

The Inputs tab indicates the units in which the metrics are stated and also gives the range of values for these variables encountered in the experimental data.

As in the fuel simulator, the user is responsible for selection of input data and for their appropriateness for use with the simulator. The structure of the simulator is that it will mechanically translate input as specified into outputs, without range checks, limitations or warnings to the user. In the studies conducted for this report, considerable attention was given to keeping the inputs within the range of the experiment that was conducted and to discount predictions of performance metrics that approached or exceed their experimental limits. If the user inputs implausible or out-of-range data for engine tests, the simulator should be expected to produce implausible or out-of-range predictions for performance.

Cautions: only the cells displaying in red font (S13:AI162) should be changed by the user. Data may be entered in this area by typing values or by using a copy/paste special/values operation. The paste special/values choice is recommended so that only the data content of cells is modified, leaving cell formatting unchanged. The user should avoid inserting or deleting rows or sorting the data area, because doing so has the potential to disrupt active formulas in the sheet. It is best to perform data manipulations in an inactive sheet elsewhere.

There are five additional tabs in the engine simulators that support the calculations in various ways. Tab Fuel Model evaluates the input test fuels in vector terms. The process is identical to that seen in the fuel simulator above. Tab Base Model evaluates the predictive models for engine performance on the base hydrocarbon fuel involved. This tab combines the vector representation of fuels with the engine and dummy variables and then evaluates the predictive models of engine performance on base hydrocarbon fuels. The coefficient values and mathematical form is documented at the top of the tab. Tab EtOH model evaluates the effect of ethanol content of the fuel in changing engine performance from that estimated for the base fuel. The coefficient values and mathematical form is documented at the top of the tab. Finally, tab Vectors contains inputs for the vector representation of fuels, and tab O-P contains graphs giving comparisons between observed and predicted values for the experimental data points as an illustration of the predictive power of the models.

10. CONCLUSIONS

Fuels can be described by various combinations of chemistry, boiling points, or physical properties. Using principal components as the basis for describing a set of experimental fuels provides the highest fidelity to the original data with the fewest number of variables. Vector representation does not eliminate correlations between fuel properties and chemistries, but rather carries those correlations through subsequent analysis, after which they can be resolved into the original fuel property and chemistry values. The advantage to this approach lies in not having to make difficult and sometimes artificial decisions at the start of analysis in order to eliminate correlated variables.

Some fuel variables affect HCCI engine operation by affecting how the engine needs to be controlled. The engine control, in turn, determines how the engine performs. The best example of this is that changes in EVCA are required to control combustion phasing across a range of octane levels and sensitivity. This, in turn, changes lambda for the engine and can have an indirect effect on ISFC. Other fuel variables directly affect engine operation without affecting control settings. The best example of this is that smoke, HC, CO, and NO_x are increased with fuels of higher aromatic content and higher boiling points.

Overall, this HCCI engine generally responded well to fuels of lower octane, higher sensitivity, lower aromatics and higher olefins, with boiling points in the lower range of those evaluated. Fuels of this type produced improved ISFC, lower NO_x, and lower smoke. This engine depends on heat from a high retained exhaust fraction for ignition and therefore emphasizes MON and sensitivity aspects of octane. Other HCCI engine concepts, where ignition might be controlled by boost or high compression ratio, might benefit from different fuel properties and chemistry.

Ethanol containing fuels, up to the 30% evaluated, appear to be satisfactory for HCCI engines of this type, and control requirements are relatively unchanged. A reduction in NO_x was noted along with an increase in HC, CO, and smoke. Loss of ISFC appears mainly related to the energy content change in the fuel. It was not possible to completely resolve ethanol effects because of the small number of fuels run and because of an engine rebuild between the two phases of the program when the ethanol blends were run. The higher octane from ethanol blending does not appear to affect HCCI ignition, perhaps due to the effects of the oxygen or the resulting higher sensitivity. Since sensitivity is involved, this conclusion may also be specific to this engine type.

The engine and fuel modeling tools have been provided along with instructions for use to allow others to conduct parametric studies of fuel and engine control effects. The models are easy to use and can be used to study fuel and engine control effects. With the use of these models comes the responsibility to ensure that any studies performed fall within the range of experimental data, where the models were developed and verified.

11. RECOMMENDATIONS FOR FUTURE WORK

The results of this study should be used to plan further research in areas of interest. The current study provided a broad overview of fuel effects for one type of HCCI engine, but was not able to fully resolve fuel effects due to the limited number of fuels and experimental limitations which have been described more fully in this report. One area of possible future study would be to further resolve chemistry and boiling point differences by using blend streams from a number of refineries. The current fuels were blended using only a single example of each blending stream which differed in both chemistry and boiling points. For example, of the octane building streams (A, B, and C), the olefin stream (C) had the lowest boiling points and the aromatic stream (A) had the highest. With the current fuels, one cannot resolve how much of the benefit of a high olefin fuel was due to the lower boiling points and how much to the octane chemistry. Further studies could also be done relative to the effects of chemistry at constant (R+M)/2 in order to determine how fuels might evolve for HCCI engines without changing the current octane grades. Further studies of ethanol blending should also be done, with consideration of distillation temperature and RVP control, in order to understand how commercial ethanol fuels, blended with reformulated gasoline blending stocks, might affect HCCI. Finally, the conclusions of this report should be evaluated on a second style of HCCI engine that uses boost or high compression ratio to trigger ignition to determine if these engines respond differently to octane values and chemistry.

In all these studies, statistics should continue to be used for experimental design, analysis of data quality, and for the building models which can be used to further the study the fuel and engine effects. PCA is particularly suited for modeling the fuels in these types of studies.

12. REFERENCES

(1) Shen, Y., "Fuel Chemistry Impacts in Gasoline HCCI," Coordinating Research Council Project No. AVFL-13 Final Report, December, 2006.

(2) Wheeler, J., "Ethanol Effects on Gasoline-Like HCCI", Coordinating Research Council Project No. AVFL-13b Final Report, August, 2009.

(3) Shen, Y., King, E., Pfahl, U., Krile, R., Slone, E., Orban, J., and Wright, K., "Fuel Chemistry Impacts on Gasoline HCCI Combustion with Negative Valve Overlap and Direct Injection", SAE Paper 2007-01-4105.

(4) Szybist, J. and Bunting, G., "Chemistry Impacts in Gasoline HCCI, Literature Review", Fuels, Engine, and Emission Research Center, Oak Ridge National Laboratory, Subcontract with AVL for benefit to CRC AVFL-13, May 2006.

(5) PCR+ in Diesel Fuels and Emissions Research. McAdams, Crawford, Hadder. March 2002. ORNL/TM-2002/16.

(6) "A Vector Approach to Regression Analysis and its Application to Heavy-Duty Diesel Emissions". McAdams, Crawford, Hadder. November 2000. ORNL/TM-2000/5.

(7) "A Vector Approach to Regression Analysis and its Application to Heavy-Duty Diesel Emissions". McAdams, Crawford, Hadder. June 2000. (SAE 2000-01-1961).

APPENDIX A. EVALUATING VECTOR CHARACTERISTICS OF FUELS

A.1 Introduction

As a reference for users of the analysis presented in this report, this appendix explains the mathematical calculations required to evaluate the vector characteristics of the gasoline-range fuels examined in the CRC AFVL test program. PCA is a well known, multivariate statistical technique that is not often used in engineering studies, but is better known in the physical sciences. For a dataset consisting of observations on N variables, PCA performs a *singular value decomposition* of the X matrix of independent variables to produce a slate of N eigenvectors and N eigenvalues.

The most common form of PCA involves decomposition of the correlation matrix; the resulting eigenvectors are normalized to a vector length of unity and the eigenvalues sum to N. In this decomposition, each eigenvector expresses one way in which the N variables are related to each other, and the eigenvalues measure how much of the variation in the dataset is associated with each eigenvector. Essentially, the correlations among the N variables are broken down into N differing patterns of variation that make up the overall relationship of variables in the dataset. PCA works on the basis of variables standardized to a mean of zero and standard deviation of one. If there are N variables, then the standardized variance in the dataset is also N.

When PCA is applied to an orthogonal dataset, the resulting eigenvectors reveal the experimental design. When PCA is applied to a non-orthogonal dataset, the eigenvectors reveal information on how the data were generated (or sampled). The eigenvectors additionally provide a set of vector variables in which the data are orthogonal. PCA has been applied extensively to the analysis of diesel engine emissions by the authors and has proved to offer reliable insight into the underlying structures of fuel datasets.

A.2 Application to AVFL Fuels

The use of PCA begins with the selection of variables that make up the fuels vector. As described in Section 4.2.3 of the report and documented in Table 4-4, a total of nine variables are required to specify a base hydrocarbon fuel:

- Six variables giving the chemical composition of the fuel according to aggregate classes: n-paraffins (excluding C_4), n-C_4, iso-paraffins, cyclo-paraffins, olefins, and aromatics.
- Three variables giving the height and shape of the distillation curve: T10, T50, and T90.

The six composition variables define the complete fuel chemistry and necessarily sum to 100%. There are only five degrees of freedom (five independent variables) among them.

The first step in computing fuel vectors begins with the transformation of the six chemistry variables to implement the physical constraint that they are bounded to values between zero and one (inclusive). Section 4.2.2 described the PCA process as forming a new coordinate system based on the characteristics seen in the data. An analogy was drawn between the cloud of data points and a football shape to explain how the coordinate system could be visualized. PCA treats

the variables making up the vectors as being normally distributed on the interval from -infinity to +infinity. The fact that the six chemistry variables sum to 100% is not a problem, because PCA can eliminate the linear dependence of the sixth variable on the other five, but the existence of physical constraints is a problem. Whenever the data cloud approaches a plane in N-space where a chemistry variable reaches a value of 0 or 1, the football shape will be truncated (has a flat spot). PCA will not recognize the constraint to the distribution of values, but rather will assume continuity of the variable across the constraining planes into non-physical territory.

To solve this problem, a mathematical transformation based on the logistic function is applied to the chemistry variables before they are entered into the PCA analysis; this process is illustrated in Figure A-1. The logistic curve – the s-shaped curve seen in the rightmost panel – is commonly used to represent processes that transition from one level to another over a finite interval. As seen in the rightmost panel, its x values are defined over the range –infinity to +infinity, but its y values are constrained to the range 0.00 to 1.00. This is very much the behavior that is wanted for the treatment of chemistry variables by PCA; in fact, it is the inverse logistic transform that is needed.

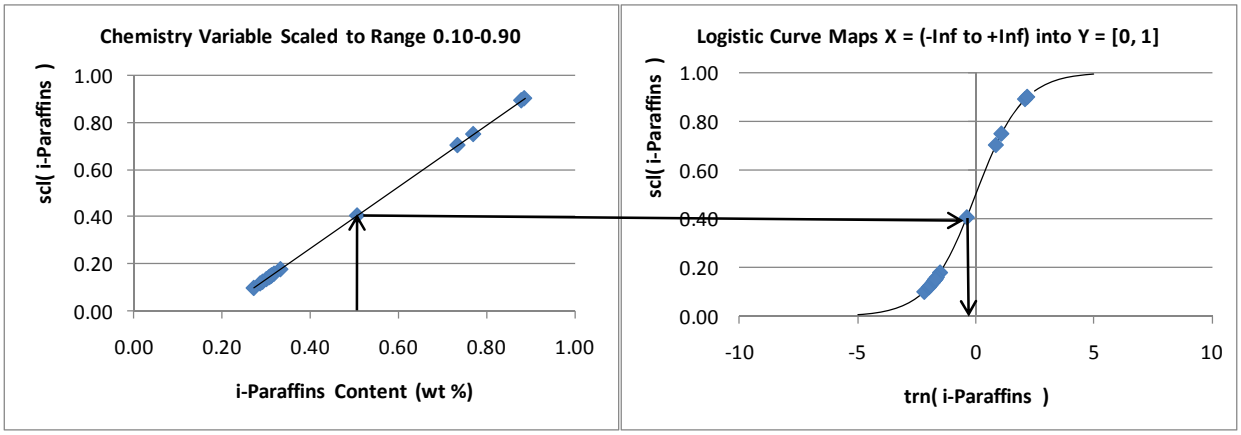


Figure A-1: Logistic Transformation of Chemistry Variables.

The transformation process for any chemistry variable is:

- Linearly scale the variable from its observed range of values to the interval 0.10 to 0.90. The smallest value occurring in the dataset is mapped to 0.10 and the largest value is mapped to 0.90. On the logistic curve, the y-axis values of 0.10 and 0.90 bound the portion of the curve that is highly linear. The scaled values of the chemistry variables become y-axis values on the logistic curve.
- Compute an x-axis value for the logistic function that is consistent with the scaled y values for each fuel.

Table A-1 gives the formulas and coefficients values needed to scale and transform the six chemistry variables used here. Table A-2 gives a worked example using Fuel A100.

Table A-1: Formulas and Coefficients Used to Compute Scaled and Transformed Values for Chemistry Variables Used in the Fuel Model.

scaled(Var) = 0.10 + 0.80· (Measured – minValue) / (maxValue – minValue)
trn(Var) = ln(scaled(Var) / (1-scaled(Var))

	Measured	minValue	maxValue
nParaXC4	As Reported (volume %)	0.0043	0.1960
nParaC4		0.0004	0.0670
iPara		0.2722	0.8865
cycloPara		0.0110	0.1693
Olefins		0.0033	0.4470
Aromatics		0.0280	0.5495

Table A-2: Worked Example for Fuel A100.

	Reported Values	Scaled Chemistry Values	Transformed Chemistry Values	Variable Value
nParaXC_4	0.1016	0.506	0.024	0.024
nParaC_4	0.0229	0.370	-0.531	-0.531
iPara	0.2722	0.100	-2.197	-2.197
cycloPara	0.0329	0.211	-1.321	-1.321
Olefins	0.0210	0.132	-1.884	-1.884
Aromatics	0.5495	0.900	2.197	2.197
T10	174.6	—	—	174.6
T50	266.2	—	—	266.2
T90	327.2	—	—	327.2

The next step in the calculation process is to normalize the PCA variable set to a mean value of 0 and standard deviation of 1 (also variance of 1). The mean and standard deviation values are those computed from the 15 experimental fuels on which the vector fuel model is based. The values required for this normalization are given in Table A-3, along with a worked example for fuel A100.

Table A-3: Normalization to Mean 0 and Standard Deviation 1, with Worked Example for Fuel A100.

	Mean Value	Standard Deviation	Variable Value	Normalized Value
trn(nParaXC$_4$)	0.096	0.057	0.024	-1.256
trn(nParaC$_4$)	0.038	0.022	-0.531	-25.974
trn(iPara)	0.455	0.235	-2.197	-11.264
trn(cycloPara)	0.095	0.049	-1.321	-28.623
trn(Olefins)	0.114	0.155	-1.884	-12.872
trn(Aromatics)	0.201	0.160	2.197	12.447
T10	174.6	18.1	158.9	0.734
T50	266.2	31.2	224.1	1.035
T90	327.2	32.4	337.1	0.213

The last step in the process is to compute the fuel's score for each of the vectors. In general, any fuel may be represented as a weighted sum of the eigenvectors:

$$F_i = s_{i,1}\cdot V_1 + s_{i,2}\cdot V_2 + s_{i,3}\cdot V_3 + s_{i,4}\cdot V_4 + \ldots + s_{i,n}\cdot V_n \qquad \text{(Eq. A-1)}$$

where the coefficients $s_{i,j}$ are the vector scores that define the extent to which fuel i expresses the characteristics (or pattern) represented by vector j. If there are N variables and N vectors in the complete fuel space, then a fuel i is completely defined by the set of N scores $\{ s_{i,j} \}$. The mean fuel in the dataset is (by definition) the fuel whose scores are identically zero. All N vectors must be used if one is to replicate the original data exactly, but for the AVFL fuels dataset only four vectors (out of 9) are required to describe the systematic variation among fuels.

The score $s_{i,j}$ for any vector j is computed simply by multiplying the fuel's description in terms of the normalized variables by the internal coefficients of the vector. The products are summed within each vector to give the final score. Tables A-4, A-5, A-6, and A-7 show worked examples for the calculation of scores for fuel A100.

Table A-4: Computation of Vector 1 Score for Fuel A100.

	Normalized Value		Vector 1 Coefficients		Vector 1 Score
trn(nParaXC$_4$)	-1.256	*	0.269	=	0.031
trn(nParaC$_4$)	-25.974	*	0.372	=	-0.214
trn(iPara)	-11.264	*	0.020	=	-0.018
trn(cycloPara)	-28.623	*	0.034	=	-0.039
trn(Olefins)	-12.872	*	-0.422	=	0.241
trn(Aromatics)	12.447	*	0.208	=	0.476
T10	0.734	*	0.445	=	0.327
T50	1.035	*	0.448	=	0.464
T90	0.213	*	0.410	=	0.087
Vector 1 Score =					**1.356**

Table A-5: Computation of Vector 2 Score for Fuel A100.

	Normalized Value		Vector 2 Coefficients		Vector 2 Score
trn(nParaXC$_4$)	-1.256	*	0.427	=	0.049
trn(nParaC$_4$)	-25.974	*	-0.246	=	0.141
trn(iPara)	-11.264	*	-0.552	=	0.486
trn(cycloPara)	-28.623	*	0.387	=	-0.449
trn(Olefins)	-12.872	*	0.253	=	-0.145
trn(Aromatics)	12.447	*	0.376	=	0.861
T10	0.734	*	0.059	=	0.043
T50	1.035	*	0.178	=	0.184
T90	0.213	*	-0.251	=	-0.054
Vector 2 Score =					**1.118**

Table A-6: Computation of Vector 3 Score for Fuel A100.

	Normalized Value		Vector 3 Coefficients		Vector 3 Score
trn(nParaXC$_4$)	-1.256	*	0.265	=	0.031
trn(nParaC$_4$)	-25.974	*	0.393	=	-0.226
trn(iPara)	-11.264	*	0.098	=	-0.086
trn(cycloPara)	-28.623	*	0.654	=	-0.758
trn(Olefins)	-12.872	*	0.092	=	-0.053
trn(Aromatics)	12.447	*	-0.544	=	-1.246
T10	0.734	*	-0.156	=	-0.115
T50	1.035	*	-0.083	=	-0.086
T90	0.213	*	0.042	=	0.009
Vector 3 Score =					**2.529**

Table A-7: Computation of Vector 4 Score for Fuel A100.

	Normalized Value		Vector 4 Coefficients		Vector 4 Score
trn(nParaXC$_4$)	-1.256	*	-0.290	=	-0.033
trn(nParaC$_4$)	-25.974	*	0.030	=	-0.017
trn(iPara)	-11.264	*	-0.143	=	0.126
trn(cycloPara)	-28.623	*	0.115	=	-0.133
trn(Olefins)	-12.872	*	0.175	=	-0.100
trn(Aromatics)	12.447	*	0.189	=	0.433
T10	0.734	*	-0.631	=	-0.463
T50	1.035	*	0.360	=	0.373
T90	0.213	*	0.536	=	0.114
Vector 4 Score =					**0.298**