United States
Environmental Protection
Agency

Air and Radiation

EPA420-P-02-001
October 2002

A Comprehensive Analysis of Biodiesel Impacts on Exhaust Emissions

Draft Technical Report

EPA420-P-02-001
October 2002

A Comprehensive Analysis of Biodiesel Impacts on Exhaust Emissions

Draft Technical Report

Assessment and Standards Division
Office of Transportation and Air Quality
U.S. Environmental Protection Agency

NOTICE

This technical report does not necessarily represent final EPA decisions or positions. It is intended to present technical analysis of issues using data that are currently available. The purpose in the release of such reports is to facilitate the exchange of technical information and to inform the public of technical developments which may form the basis for a final EPA decision, position, or regulatory action.

Nature and Purpose of This Technical Report

This Report presents a technical analysis of the effect of biodiesel on exhaust emissions from diesel-powered vehicles. It analyzes pre-existing data from various emissions test programs to investigate these effects. The conclusions drawn in this Technical Report represent the current understanding of this specific technical issue, and are subject to re-evaluation at any time.

The purpose of this Technical Report is to provide information to interested parties who may be evaluating the value, effectiveness, and appropriateness of the use of biodiesel. This Report informs any interested party as to the potential air emission impacts of biodiesel. It is being provided to the public in draft form so that interested parties will have an opportunity to review the methodology, assumptions, and conclusions. The Agency will also be requesting independent peer reviews on this draft Technical Report from experts outside the Agency.

This Technical Report is not a rulemaking, and does not establish any legal rights or obligations for any party. It is not intended to act as a model rule for any State or other party. This Report is by its nature limited to the technical analysis included, and is not designed to address the wide variety of additional factors that could be considered by a State when initiating a fuel control rulemaking. For example, this Report does not consider isues such as air quality need, cost, cost effectiveness, technical feasibility, fuel distribution and supply impacts, regional fleet composition, and other potentially relevant factors.

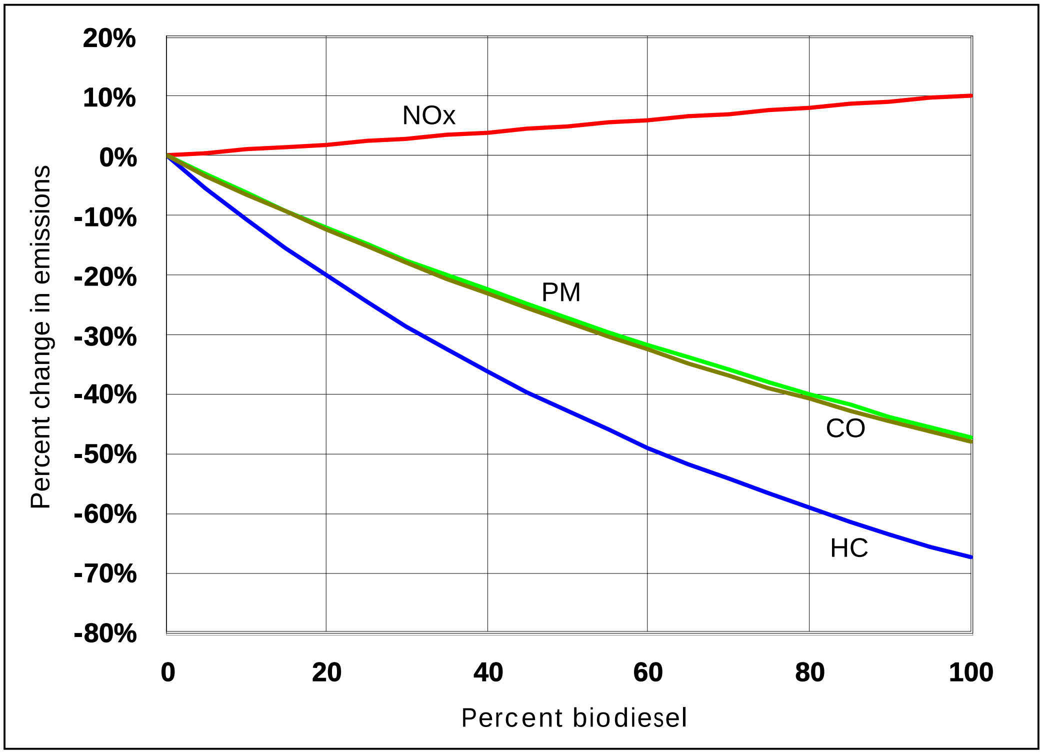
State or local controls on motor vehicle fuels are limited under the Clean Air Act (CAA) - certain state fuel controls are prohibited under the Clean Air Act, for example where the state control applies to a fuel characteristic or component that EPA has regulated (see CAA Section 211(c)(4)). This prohibition is waived if EPA approves the State fuel control into the State Implementation Plan (SIP). EPA has issued guidance describing the criteria for SIP approval of an otherwise preempted fuel control. See “Guidance on the Use of Opt-in to RFG and Low RVP Requirements in Ozone SIPs,” (August, 1997) at: http://www.epa.gov/otaq/volatility.htm.

The SIP approval process, a notice and comment rulemaking, would also consider a variety of technical and other issues in determining whether to approve the State fuel control and what emissions credits to allow. An EPA Technical Report like this one can be of value in such a rulemaking, but the SIP rulemaking would need to consider a variety of factors specific to the area, such as fleet make-up, refueling patterns, program enforcement and any other relevant factors. Additional evidence on emissions effects that might be available could also be considered. The determination of emissions credits would be made when the SIP rulemaking is concluded, after considering all relevant information. While a Technical Report such as this may be a factor in such a rulemaking, the Technical Report is not intended to be a determination of SIP credits for a State fuel program.

Executive Summary

Due to the increasing interest in the use of biodiesel, the Environmental Protection Agency has conducted a comprehensive analysis of the emission impacts of biodiesel using publicly available data. This investigation made use of statistical regression analysis to correlate the concentration of biodiesel in conventional diesel fuel with changes in regulated and unregulated pollutants. Since the majority of available data was collected on heavy-duty highway engines, this data formed the basis of the analysis. The average effects are shown in Figure ES-A.

Figure ES-A
Average emission impacts of biodiesel for heavy-duty highway engines



One of the most common blends of biodiesel contains 20 volume percent biodiesel and 80 volume percent conventional diesel. For soybean-based biodiesel at this concentration, the estimated emission impacts for the current fleet are shown in Table ES-A.

Table ES-A
Emission impacts of 20 vol% biodiesel
for soybean-based biodiesel added to an average base fuel

	Percent change in emissions
NOx	+ 2.0 %
PM	- 10.1 %
HC	- 21.1 %
CO	-11.0 %

Biodiesel is also predicted to reduce fuel economy by 1-2 percent for a 20 volume percent biodiesel blend. Aggregate toxics are predicted to be reduced, but the impacts differ from one toxic compound to another. We were not able to identify an unambiguous difference in exhaust CO_2 emissions between biodiesel and conventional diesel. However, it should be noted that the CO_2 benefits commonly attributed to biodiesel are the result of the renewability of the biodiesel itself, not the comparative exhaust CO_2 emissions. An investigation into the renewability of biodiesel was beyond the scope of this report.

We have high confidence in these estimates for the current fleet. However, the database contained no engines equipped with exhaust gas recirculation (EGR), NOx adsorbers, or PM traps. In addition, approximately 98% of the data was collected on 1997 or earlier model year engines. We made an attempt to estimate the impacts that biodiesel might have on EGR-equipped engines by investigating cetane effects of biodiesel, and we have no reason to believe that biodiesel will have substantially different impacts on emissions from the effects shown above for engines having NOx adsorbers or PM traps. Still, our estimates of biodiesel impacts on emissions may be less accurate for future fleets than they are for the current fleet.

The investigation also discovered that biodiesel impacts on emissions varied depending on the type of biodiesel (soybean, rapeseed, or animal fats) and on the type of conventional diesel to which the biodiesel was added. With one minor exception, emission impacts of biodiesel did not appear to differ by engine model year.

The highway engine-based correlations between biodiesel concentration and emissions were also compared to data collected on nonroad engines and light-duty vehicles. On the basis of this comparison, we could not say with confidence that either of these groups responded to biodiesel in the same way that heavy-duty highway engines do. Thus we cannot make any predictions concerning the impacts of biodiesel use on emissions from light-duty diesel vehicles or diesel-powered nonroad equipment.

Table of Contents

Section I: Introduction

There has been increasing interest in recent years in the use of biodiesel as a substitute for petroleum-based diesel fuel. This draft technical report describes an assessment of the effect of biodiesel on exhaust emissions of regulated and unregulated pollutants for vehicles and engines that have not been specifically modified to use biodiesel. This draft report is intended as a starting place for discussion and comment. By analyzing the emission impacts of biodiesel, we have drawn no conclusions regarding the appropriateness of its use for any particular purpose or in any particular context. Rather, this technical assessment of emissions impacts is intended only to inform parties that are considering the use of biodiesel.

A. Regulatory Context

As States review their programmatic options for meeting air quality goals, biodiesel is being considered more frequently. Several municipalities and States are considering mandating the use of low levels of biodiesel in diesel fuel on the basis of several studies which have found hydrocarbon (HC) and particulate matter (PM) benefits from the use of biodiesel. Biodiesel may be appealing for other reasons as well. Renewed concern about national energy security has heightened interest in the use of biodiesel as a domestically-produced diesel fuel substitute. There is also strong evidence that biodiesel can reduce emissions of greenhouse gases, particularly when emissions generated during its full production-to-consumption lifecycle are taken into account.

Unfortunately, the studies which have examined biodiesel emission effects have not been entirely consistent in their conclusions, and some studies also suggest that the use of biodiesel may produce small increases in emissions of oxides of nitrogen (NOx) concurrent with reductions in other pollutants. For States wishing to account for any potential air quality benefits of biodiesel use, this presents a dilemma for U.S. EPA reviewers of State Implementation Plans. As a result of the substantial recent interest in biodiesel and the lack of comprehensive information on its emission impacts, in August 2001, the Environmental Protection Agency (EPA) initiated an effort to evaluate the emission benefits of biodiesel for diesel engines which had not been specifically modified to be used with biodiesel.

B. Objectives and Scope of Research

The primary goal of this EPA project is to provide an objective estimate of the effect of biodiesel use on emissions of regulated and unregulated pollutants using existing data. As such, our objective is to provide correlations between the concentration of biodiesel in conventional diesel fuel and the percent change in emitted levels of different categories of pollutants, as well as fuel economy. The program is not intended to include investigations of other aspects of biodiesel use, such as:

Engine durability
Materials compatibility
Fuel storage stability
Lubricity
Renewability/full fuel lifecycle emissions
Biodiesel production feedstocks or costs
Cold flow properties
Cost

As a result, the tentative conclusions reached in this report have some limitations. For instance, the data on which we based our analyses was collected before the biodiesel samples had an opportunity to degrade due to excessive storage time. Therefore, our estimates of emissions impacts for biodiesel use would not be applicable to cases in which the biodiesel has experienced some degradation. Likewise our estimates would not apply to situations in which cold flow may be a problem. We also make no claims in this technical report regarding the comparative health effects of PM emissions from biodiesel versus PM emissions from conventional diesel, other than to estimate changes in total PM mass emissions.

We did investigate the degree to which other factors influenced the relationship between biodiesel and emissions, including:

Engine/vehicle technology
Highway versus nonroad engines
Test cycle
Base fuel to which biodiesel is added
Light versus heavy-duty
Type of biodiesel (soybean, rapeseed, grease)

Data on fuel economy, carbon dioxide (CO_2) emissions, and toxics were much more limited than the data on NOx, HC, PM, and carbon monoxide (CO). In particular, correlating biodiesel with toxics required a more limited statistical approach than that used to evaluate the other pollutants, and did not permit us to investigate many of the additional factors described above. Data on light-duty vehicles and nonroad engines was also quite limited. Therefore, we used data on light-duty and nonroad as validation sets to determine if correlations based on heavy-duty highway engines could be used as an appropriate predictor of emissions from light-duty or nonroad fueled with biodiesel.

C. Interaction with Stakeholders

To assure that our analysis benefitted from data and expertise outside the EPA, we have made a deliberate effort to work with external stakeholders from the beginning of our work on this topic. In order to inform stakeholders of the work that EPA would be doing to look at the effects of biodiesel on emissions, informational letters were sent to stakeholders. We also established an e-mail list for individuals known to have an interest in our work, with periodic messages sent to inform members of the status of our work. We also developed a web site (http://www.epa.gov/otaq/models/biodsl.htm) to share our plans and intermediate work products.

Comments on this draft technical report and our analysis should be sent by December 31, 2002 to David Korotney at korotney.david@epa.gov, or through regular mail to:

David Korotney
U.S. EPA National Vehicle and Fuel Emissions Laboratory
2000 Traverwood Drive
Ann Arbor, MI 48105

To assure that our correlations represent the best current scientific understanding of the emission impacts of biodiesel, we also intend to conduct a workshop subsequent to the comment period on this draft technical report to discuss technical issues related to our analysis. For information on this workshop, please see our website. At the conclusion of the workshop, EPA will consider all the comments received and will revise our analysis in response to those comments. We then plan to publish a final technical report that summarizes our work and conclusions.

Section II: What Data Was Used?

We began the process of assembling data for use in correlating biodiesel concentrations with emissions by conducting literature searches and reviewing lists of relevant data sources that had been assembled by other researchers for use in similar analyses. Once we had assembled a complete list of prospective data sources, we reviewed each study to verify that it contained the actual raw data that the report or study described. If the raw data was not provided, we made attempts to contact the authors. The complete list of data sources that we considered for our analysis is given in Appendix A. Studies that were excluded from our analysis are separated and categorized according to the reason for their exclusion.

We reviewed the studies to verify that they met certain criteria consistent with the goals of the project. These criteria are described in Section II.A below. As a result of this review, only 39 of the full set of 80 studies were retained for our analysis. We then entered the data into a database specifically designed for this project, making adjustments to ensure consistency in units and corrections for emissions drift over time. All of these steps are described in the remaining portions of this Section.

A. Criteria for choosing data sources

The data that we considered for use in this analysis was screened to ensure that it met certain criteria. For instance, we limited our analysis to No. 1 and No. 2 diesel fuel and related blends that can be used in a typical heavy-duty diesel engine without engine modifications. As a result we excluded all emulsions and non-biodiesel oxygenated blends with more than 20 vol% oxygenate. We also excluded fuels that were made entirely from pure chemicals rather than refinery streams. We did not specifically exclude Fischer-Tropsch fuels, nor did we limited ourselves to diesel fuels containing less than 500 ppm sulfur.

We also limited this study to engines that had already been sold commercially or had a high probability of being sold in the future. Engines with experimental technologies that had no immediate plans for commercialization, such as those with innovative combustion chamber geometries, were excluded. Likewise, single-cylinder research engines were also excluded from consideration even though the associated full-size parent engine might have been appropriately included in the database had it been tested. Single-cylinder engines do not appear in heavy-duty applications. By definition they have lower total horsepower and displacement, both of which may influence the way in which biodiesel impacts emissions. Unless we were to make the assumption that single cylinder engines respond in the same way as their parent engine to changes in fuel properties, we would have to define new technology groups specific to single-cylinder engines. Light-duty vehicle and nonroad data was not specifically excluded from the analysis, but the paucity of this data made it necessary for us to evaluate their effects separately.

We also excluded data that was collected under test cycles that were unique or in some way unrepresentative of the Federal Test Procedure (FTP). For instance, a number of studies tested an engine only under a single steady-state mode, while others used two or three nonstandard modes for testing. However, there were a number of studies that used atypical test cycles which were comprehensive enough in their number and/or selection of modes, or in the design of their transient speed-load traces, that the resulting emission measurements may still be informative. These latter observations were identified in the database as having the generic test cycle label TRANSIENT or STEADY-STATE as applicable, and were analyzed separately from the rest of the data.

The type of testing also played a role in determining if a given study should be included in our analysis. For instance, since we were primarily interested in biodiesel impacts on emissions, we excluded all studies that did not test at least two different fuels on the same engine at two different biodiesel concentrations (one of which could be 0% biodiesel). Also, we considered only studies in which the base fuel to which a biodiesel blend was compared was the same conventional diesel fuel used to create the biodiesel/diesel blend.

There were a number of cases in which data from one study was repeated in another study. This might occur if the authors published the same data set in multiple scientific journals to maximize exposure, or if the authors presented a previously-published set of data in a new publication for the purpose of comparing the two datasets. Table II.A-1 lists the cases in which repeat publications were excluded from our database.

Table II.A-1
Exclusion of duplicate datasets from the database

Retained Study	Excluded Study
Graboski, M.S., J.D. Ross, R.L. McCormick, "Transient Emissions from No. 2 Diesel and Biodiesel Blends in a DDC Series 60 Engine," SAE paper no. 961166	Colorado Institute for Fuels and High Altitude Engine Research, "Emissions from Biodiesel Blends and Neat Biodiesel from a 1991 Model Series 60 Engine Operating at High Altitude," Final Report to National Renewable Energy Laboratory, September 1994 Note: CO_2 values were not duplicative
Peterson, C.L., "Truck-In-The-Park Biodiesel Demonstration with Yellowstone National Park," University of Idaho, August 1999.	Peterson, C.L., D.L. Reece, "Emissions Testing with Blends of Esters of Rapeseed Oil Fuel With and Without a Catalytic Converter," SAE paper no. 961114 Note: Only Table 6 data is duplicative
Manicom, B., C. Green, W. Goetz, "Methyl Soyate Evaluation of Various Diesel Blends in a DDC 6V-92 TA Engine," Ortech International, April 21, 1993	Schumacher, L.G., S.C. Borgelt, W.G. Hires, D. Fosseen, W. Goetz, "Fueling Diesel Engines with Blends of Methyl Ester Soybean Oil and Diesel Fuel," University of Missouri

Sharp, C.A., S.A. Howell, J. Jobe, "The Effect of Biodiesel Fuels on transient Emissions from Modern Diesel Engines, Part I Regulated Emissions and Performance," SAE paper no. 2000-01-1967	Sharp, C.A., "Characterization of Biodiesel Exhaust Emissions for EPA 211(b)," Final Report on Cummins N14 Engine, prepared for National Biodiesel Board, January 1998
Peterson, C.L., "Truck-In-The-Park Biodiesel Demonstration with Yellowstone National Park," University of Idaho, August 1999.	Taberski, J.S., C.L. Peterson, "Dynamometer Emissions test Comparisons on a 5.9L Direct Injected Diesel Powered Pickup," *BioEnergy '98: Expanding BioEnergy Partnerships* Taberski, J.S., C.L. Peterson, J. Thompson, H. Haines, "Using Biodiesel in Yellowstone National Park - Final Report of the Truck in the Park Project," SAE paper no. 1999-01-2798
Graboski, M.S., R.L. McCormick, T.L. Alleman, A.M. Herring, "The Effect of Biodiesel Composition on Engine Emissions from a DDC Series 60 Diesel Engine," Colorado School of Mines, Final Report to National Renewable Energy Laboratory, June 8, 2000	McCormick, R.L., M.S. Graboski, T.L. Alleman, A.M. Herring, "Impact of Biodiesel Source Material and Chemical Structure on Emissions of Criteria Pollutants from a Heavy-Duty Engine," *Environmental Science and Technology*, **2001**, *35*, 1742-1747

B. Preparation of database

1. Database structure

In designing the structure of the database and the fields that would be included, it was our intention to include all information that had any potential for helping us to quantify the relationship between biodiesel concentration and emissions. In addition, we also wanted to ensure that a wide variety of issues could be investigated once the database was assembled, including issues which were not immediately germane to our primary goal of correlating biodiesel concentration with emissions. This secondary goal is of broad and continuing interest to the EPA as we continue our efforts to understand and control pollution from diesel-powered engines and vehicles. Towards these ends, we selected a wide variety of fuel, engine, and test parameters to include in the database.

The database was divided into three separate files:

Fuels.xls	File containing a complete description of every fuel, including physical, compositional, and chemical characteristics.
Equipment.xls	File containing a complete description of every engine, include both engine design characteristics and elements that may have been changed subsequent to production, such as aftertreatment and EGR
Emissions.xls	File containing individual test descriptions and emission results

Data source IDs were used to link specific fuels, engines, and emission estimates across the three files. A complete description of the fields for all three database files is given in Appendix B.

2. Entering data

The primary concern as data was being entered into the database was consistency of units. For the most part, these conversions were straightforward. In some cases, however, the fuel property unit conversions were not straightforward due to ambiguity in either a given study or the database structure itself. In these cases, decisions were made that were intended to maximize the useful amount of data. These decisions are summarized below:

> *Viscosity* - The viscosity of a fuel can be measured at different temperatures. In cases where more than one temperature was used, the measurement closest to 40 ^{o}C was entered into the database. If only one viscosity measurement was made, it was entered into the database without regard to test temperature.
>
> *Oxygen* - If an oxygenate was not added to a fuel and the oxygen level was not measured, it was assumed to be zero. If oxygen was measured, we used the measured value even if doing so included the oxygen contribution of, for instance, cetane improver additives.
>
> *Properties of cetane-enhanced fuels* - If the properties of a fuel were measured before a cetane improver was added to the fuel but not afterward, the properties of the base fuel were considered to be applicable to the additized fuel as well, with the exception of cetane number.
>
> *Concentration of cetane improver additives* - Our database required that the concentration of cetane improver additives be entered as vol%. If a study provided the concentrations in terms of wt%, the conversions were made using the following equation:
>
> $$\text{vol\%} = \text{wt\%} \text{ X fuel specific gravity} / b$$
>
> where b is the specific gravity of the cetane improver additive[a].
>
> *Cetane increase due to additives* - If the increase in cetane number which resulted from the addition of a cetane improver additive was not given in the study, it was estimated from a correlation given in SAE paper number 972901. This correlation is:
>
> $$CNI = a \times CN^{0.36} \times G^{0.57} \times C^{0.032} \times \ln(1 + 17.5 \times C)$$

---

[a] For 2-ethylhexylnitrate (EHN) this value is 0.964 according to an Ethyl data sheet on their HiTEC Cetane Improver Additive (composed of 99% 2-EHN). For di-tertiary butylperoxide (DTBP), the value of b is 0.794 according to the CRC Handbook of Chemistry and Physics.

Where:

CNI = Predicted cetane number increase due to an additive
a = 0.16 for 2-ethylhexylnitrate and 0.119 for di-tertiary butylperoxide
CN = Base cetane number
G = Fuel API gravity
C = concentration of additive in vol%

Cetane index - If the cetane number of a fuel was not measured, the cetane index was used to estimate the cetane number. An analysis of unadditized fuels in the survey database collected by the Alliance of Automobile Manufacturers indicates that cetane index does not have a 1:1 correlation with natural cetane number as formerly believed. Instead, the following equation appears to provide a much more precise relationship for fuels in which no cetane improver additives were used:

$$\text{Natural cetane number} = 1.154 \times \text{Cetane index} - 9.231$$

This equation was used to estimate the natural cetane number is cases where only the cetane index was given and the fuel contained no cetane improver additives.

Aromatics test methods - The database required total aromatics content to be entered in units of vol% as established from an FIA test method (ASTM D 1319 or the equivalent). If total aromatics content was derived using supercritical fluid chromatography (SFC, from ASTM D 5186 or its equivalent), which produces measurements in wt%, the conversion was made using an equation derived from the California Code of Regulations, Title 13, Section 2282(c)(1):

$$\text{vol\% (by FIA)} = 0.916 \times \text{wt\% (by SFC)} + 1.33$$

If total aromatics content was not measured by an SFC test method, then alternative conversion equations were used. These conversion equations are described in Appendix E.

Total, mono, and polyaromatics - Total aromatics content is the sum of mono and polyaromatics. Thus if a study provided measurements for only two of these three properties, the third was estimated based on this relationship. Mono and polyaromatics was entered as weight percent.

The database required mono and polyaromatics to be entered in units of wt% as established using an SFC test method. If total aromatics was not measured by an SFC test method, then the conversion to wt% by SFC was made using equations described in Appendix C of the July 2001 Staff Discussion Document.

There were also situations in which aspects of the data not related to fuel properties were ambiguous. In these situations, we again made decisions that were intended to maximize the usefulness of the database in the context of developing correlations between biodiesel concentration and emissions. The primary decisions are listed below:

Hot-start versus composite FTP - If the heavy-duty transient Federal Test Procedure was used to produce composite emission measurements, these were labeled as "UDDS" cycle values in the database (for the Urban Driving Dynamometer Schedule, the schedule on which the FTP is based). If the FTP was used to produce separate hot and cold-start emission measurements and no composite results were presented, then the hot and cold-start results were weighted at 6/7 and 1/7, respectively, to produce composite results which were then entered into the database as UDDS cycle values. In this process, all available hot-start tests were averaged before calculating the composite value. If the FTP was used to produce only hot-start emission measurements, then these results were entered into the database as UDDSH cycle values.

Engine adjustments - If adjustments were made to an engine (such as changes in injection timing, addition or removal of aftertreatment, etc.), these were treated as unique engines and entered into the database as such. Thus each engine value in the database refers to a set of emissions data from a single engine whose operating parameters and physical characteristics did not change during the course of testing.

Repeat measurements - There were many cases in which the same fuel was tested on the same engine multiple times. All such repeat measurements were entered into the database.

Averaged emissions - If the study presented only averaged emissions resulting from multiple repeat tests of a single fuel on a single engine, the average values were entered into the database the same number of times as the number of repeat tests on which the average was based. If the number of repeat tests was unknown, it was assumed to be two.

3. Adjustments to database

Once all the data had been entered into the database and it had been reviewed for errors and inconsistencies, some adjustments were made to ensure that the database was best suited for our analysis.

The first adjustment involved correcting for engine drift over time. This correction was necessary for cases in which the emissions from an engine appeared to drift upwards or downwards over the course of the study. Cases in which this drift was evident were those in which the study authors specifically looked for it by testing a single fuel - usually a reference fuel - at multiple times throughout the test program. The emissions from this fuel could then be plotted against time (engine hours, date, or test run number) to determine if drift occurred. If engine drift was evident, the authors may have chosen to add a time parameter to the regression equations that were developed using the study data instead of correcting the data itself. This option was not available in our correlation because so few studies included time measurements. Thus it was necessary for us to correct the data for those studies in which engine drift was investigated and found to be significant.

For the particular studies in our database, we found it was not necessary to use the reference fuel in a given study to generate a correlation between emissions and time. Instead, some studies included adjusted data that had already been corrected for time drift. For other studies, time drift appeared more as a step change than as a continuous function. In these latter cases it was possible to divide all the reference fuel emission measurements into independent groups, and then use each group as the reference for biodiesel emission measurements collected in the same timeframe. Table II.B.3-1 lists the studies affected and the type of correction that was made to account for engine drift.

Table II.B.3-1
Studies corrected for engine drift

Study	Engine drift correct
SAE paper no. 961166	Time-drift adjusted data provided in paper
Colorado School of Mines 1994	Time-drift adjusted data provided in paper
Fosseen 1994b	Time-drift adjusted data provided in paper
Graboski 2000	Base fuel tests were found in four different groups, each of which was statistically distinct from the others. These four groups were used as separate reference fuels in lieu of implementing a time-drift correction equation
McCormick 2001	Base fuel tests after Feb. 16, 2001 were treated as a separate group to account for time drift

There were also some special cases in which data was not entered into the database in exactly the same form that it was presented in the study. For instance, if multiple hot-start measurements were taken under the FTP in a given study, but only one of those hot-start measurements was used to calculate the composite (at 1/7 weighting for cold-start and 6/7 weighting for hot-start), the composite value was recalculated by first averaging all available hot-start measurements. In this way all available data was used, though the composite value in the database may not be exactly the same as the composite value calculated by the study's authors. There were also cases in which we rejected cold-start data altogether. This occurred in cases where cold start measurements may have been taken on a reference fuel but not on the biodiesel blend, or vice-versa. Since the assessment of biodiesel impacts on emissions requires that all other variables remain constant when comparing a reference fuel to a biodiesel blend, we chose to exclude the cold-start data in such cases to maintain consistency in test cycle. These types of data exclusions occurred with the following two studies:

- McCormick 2001
- Graboski 2000

C. Emission standards groups

Engines with different technologies may respond differently to the use of biodiesel. It therefore seemed prudent to examine ways in which the database could be subdivided to capture any technology-specific effects that might exist. Unfortunately, the engine parameter data available in the studies that comprised our database was rather limited, and did not permit the precise assignment of individual engines to particular technology groups.

We took several alternative approaches to investigating the impact that engine technology on the correlation between biodiesel use and emissions. One approach was to categorize engines by the emission standards they were designed to meet as a surrogate for engine technology groupings. Since the model year was available for nearly every engine in the database, the engines were simply grouped by those model years that were required to meet a single set of emission standards. These model year groups are shown in Table II.C-1

Table II.C-1
Engine standards groups for heavy-duty highway diesel engines

Standards group	Model years	Federal emission standards, g/bhp-hr				
		HC	CO	NOx	HC + NOx	PM
B	2002 - 2006	-	15.5	-	2.4[α]	0.10[β]
C	1998 - 2001	1.3	15.5	4.0	-	0.10[β]
D	1994 - 1997	1.3	15.5	5.0	-	0.10[β]
E	1991 - 1993	1.3	15.5	5.0	-	0.25[β]
F	1990	1.3	15.5	6.0	-	0.60
G	1988 - 1989	1.3	15.5	10.7	-	0.60
H	1984 - 1987[γ]	1.3	15.5	10.7	-	-
I	- 1983[δ]	1.5	25	-	10	-

[α] Non-methane HC. Manufacturers have an option of meeting a 2.5 g/bhp-hr standard with a 0.5 g/bhp-hr cap on NMHC

[β] Standard for urban buses is lower.

[γ] For 1984 model years, manufacturers could opt to certify on the 13 mode steady-state cycle

[δ] Standards shown applied to 1979 - 1983 model years. However, earlier model years have been grouped with 1979 - 1983 model years for the purposes of this analysis.

Dividing the data by standards groups is also ideal from the standpoint of correlating the results of our analysis with the in-use fleet, since emission inventories are currently determined as a function of vehicle model year and age. Although engines of a given model year can have widely varying technologies, and some specific engine technologies span many model years, we believe that this approach is an appropriate alternative to technology groups. It is noteworthy that a recent analysis of the impact of diesel fuel properties on emissions from heavy-duty engines[1] found that, except for a few unique cases, engine technology does not play a significant role in the way that engines respond to changes in fuel properties.

There were several other ways that we investigated the impact of engine technology on the correlation between biodiesel and exhaust emissions. One was to exclude from our curve-fitting analyses observations that were based on technologies with potentially different biodiesel effects compared to heavy-duty highway diesel engines. This included nonroad engines and light-duty vehicles. After the curve-fitting analyses were completed, data for these excluded engines/vehicles was compared to the correlations between biodiesel use and emissions to determine if these excluded engines/vehicles responded the same way to biodiesel as heavy-duty highway engines. We also examined plots of the data to determine if certain groupings of observations could be attributed to a consistent collection of engine technologies. These analyses are described in more detail in the Sections IV and V of this report.

D. Test cycles

The studies that we reviewed contained data generated from several different test cycles. The following is a description of the various test cycles, both transient and steady-state, that we evaluated throughout this work. We also provide a description of the test cycles that were chosen for evaluating the effects of biodiesel on emissions. In selecting test cycles, we aimed at selecting those cycles that were most representative of in-use operations.

There were several transient test cycles used in the studies included in our database. By far the most prominent was the Urban Driving Dynamometer Schedule (UDDS). This test cycle forms the basis of the Federal Test Procedure (FTP) used for engine certification, and we have considered the FTP to be the most representative of in-use operation, especially for particulate emissions. The heavy-duty, on-highway FTP consists of a variety of different speeds and loads that are sequenced to simulate the urban operation running of the vehicle that corresponds to the engine being tested. The average load factor of the heavy-duty FTP cycle is roughly 20 to 25 percent of the maximum engine horsepower available at a given speed. In our database, we refer to the EPA transient test as the UDDS. The EPA transient cycle run with a hot start only is referred to as UDDSH.[b]

There were also some studies that used transient test cycles which were different than the FTP. These other transient test cycles represented only about 5 percent of all the transient data in the database. Emission measurements made under these alternative transient test cycles were identified in the database as having the generic test cycle label TRANSIENT, and were analyzed separately from the rest of the data.

The ECE R49 cycle (also called the EEC 88/77 cycle) is the 13-mode steady-state test cycle for heavy-duty diesel engines which was used for certification of heavy-duty engines in Europe until October 2000. The test cycle is similar to the US 13-mode cycle, as both cycles have identical running conditions. However, the R49 has different weighting factors at the idle speeds and is characterized by high engine loads. Nevertheless, it is considered to be an appropriate representation of some types of in-use engine operation, at least for NOx and HC. We have therefore grouped R49 data with FTP data during our analyses for these two pollutants. See Section IV.B.1 for a comparison of R49 and FTP impacts on the relationship between biodiesel concentration and emissions of regulated pollutants.

In selecting data to include in our correlations, the choice of test cycle was considered to be very important. Data generated from UDDS (FTP) transient cycle was preferred, as this cycle

---

[b] EPA's Mobile Source Observation Database (MSOD), the database of in-use vehicle test result data, uses UDDS as a test procedure value. However, this cycle is often referred to by others as a FTP test cycle. Our correlation did draw from MSOD on the data design, with a key distinction being that MSOD distinguishes between test procedures and schedules while this correlation does not.

most closely represents in-use conditions. A number of studies only measured hot-start transient emissions. When this was the case, we included this data in our database and considered this data to be satisfactory when developing our correlations between biodiesel use and emissions. As hot-start results comprise 6/7 of the composite value, we assumed that fuel effects measured using the hot-start transient test could be considered representative of composite results. We tested this assumption during our analysis and concluded that it was reasonable. See Section IV.B.1 for details.

Our decision to include certain steady-state NOx and HC emission data in the correlation is confirmed by a previous study that found that fuel modifications produce similar changes in emissions over the R49 and the heavy-duty FTP tests[2]. This study concluded that the effects of fuel property changes on emissions were similar and that general extrapolations of effects from steady-state data to transient operation are reasonable.

E. Summary statistics of data

This Section provides information on the data in our database, including distribution of fuel properties, test cycles, and model years. This information can be used to assess the degree to which the data used to develop our correlations are representative of in-use fuels and engines. The summaries in this Section include all data in the database, i.e. no outliers identified during the analysis or observations with incomplete data have been excluded in these summaries, unless specified otherwise.

1. Fuel properties

Because the analysis was intended to assess how the use of biodiesel affects emissions, we began by investigating the properties of biodiesel and comparing those properties to conventional diesel fuel. Of the 31 neat biodiesels in our database, 12 included a full complement of measured fuel properties, while nearly all included measurements for natural cetane and specific gravity. We determined average fuel properties for the neat biodiesels and compared them to average fuel properties for conventional diesel fuel sold outside of California (based on survey data from the Alliance of Automobile Manufacturers). The results are shown in Table II.E.1-1

Table II.E.1-1
Comparison between biodiesel and diesel fuel outside of California

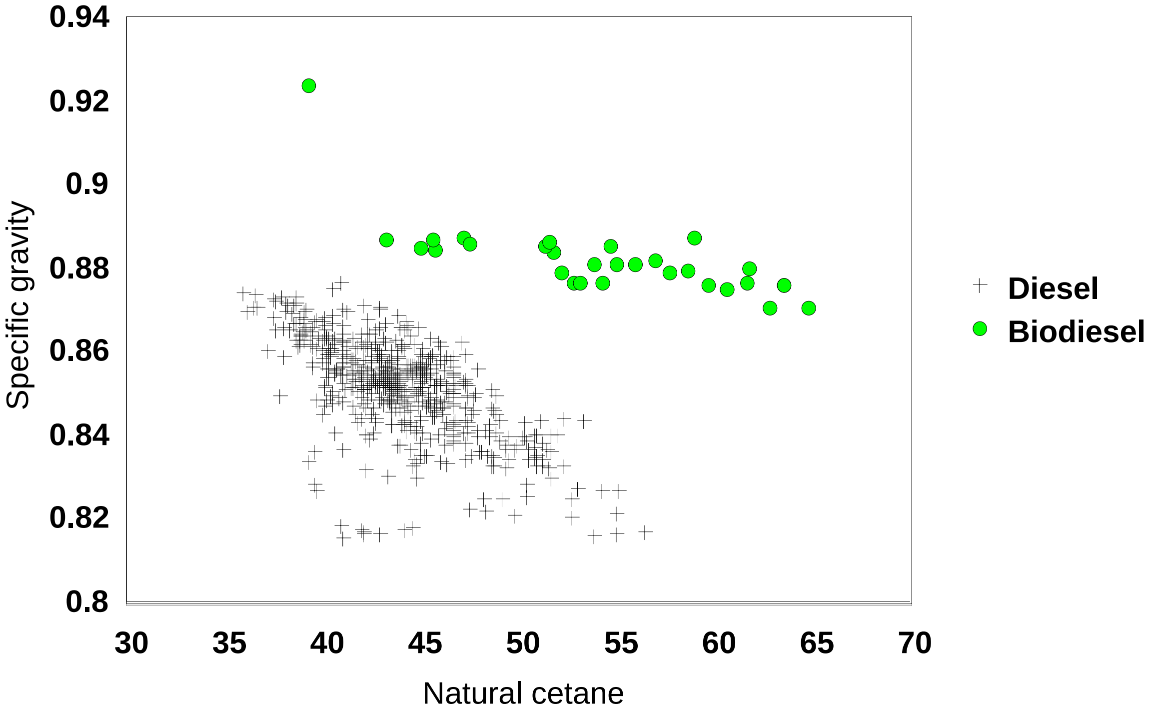
	Average biodiesel	Average diesel
Natural cetane number	55	44
Sulfur, ppm	54	333
Nitrogen, ppm	18	114
Aromatics, vol%	0	34
T10, °F	628	422
T50, °F	649	505
T90, °F	666	603
Specific gravity	0.88	0.85
Viscosity, cSt at 40 °F	6.0	2.6

The neat biodiesels used to calculate the above average values can be subdivided into several broad groups. These include virgin oils versus their transesterified counterparts, and plant versus animal-based biodiesels. The plant-based biodiesels in the database are derived from soybean, rapeseed, and canola oils, while the animal-based biodiesels are derived from tallow, grease, and lard. A more detailed discussion of how we subdivided the plant and animal-based biodiesels can be found in Section III.C.2.c.

The largest group is the plant-based esters, comprising nearly 80% of all the biodiesel blends in the database. Animal-based esters comprise most of the remaining biodiesel blends. The database contains only two virgin oils, and these appear to have significantly different fuel properties from the esters. As a result we removed the virgin oil biodiesels from our curve-fitting, and analyzed them separately (see Section IV.D).

To illustrate how biodiesel and conventional diesel differ, we also examined the distribution of specific gravity and natural cetane, since these two fuel properties were measured for nearly every one of the neat biodiesels in the database. Figure II.E.1-1 shows the results.

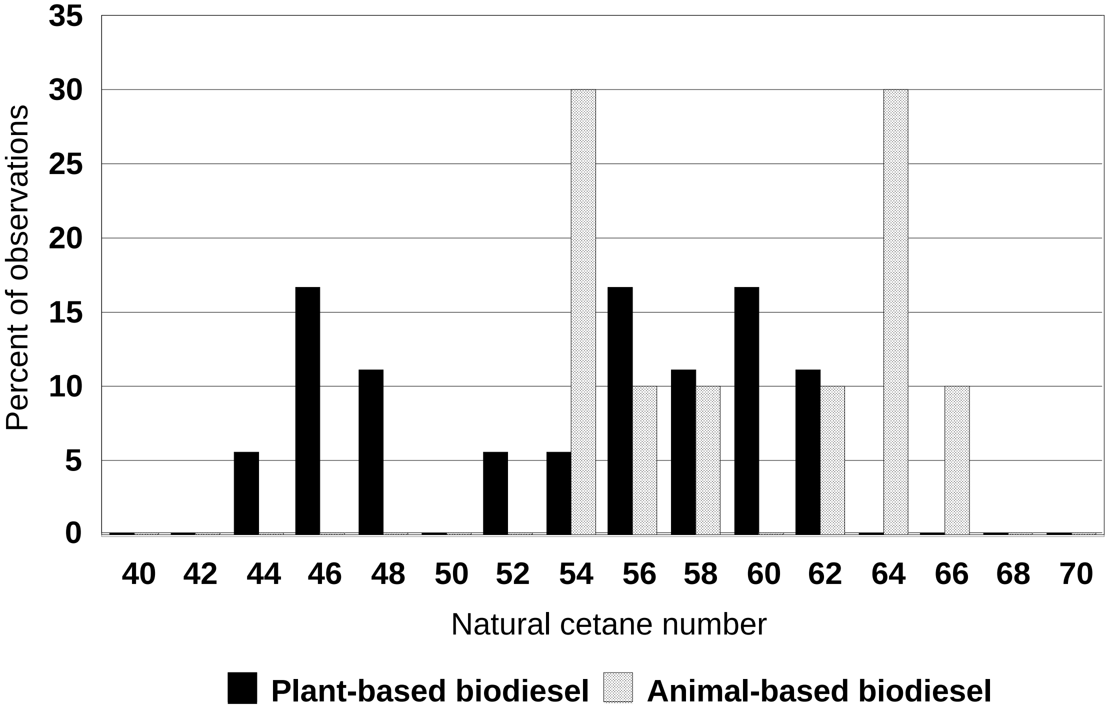
Figure II.E.1-1
Natural cetane and specific gravity of biodiesel and conventional diesel



In this figure, the biodiesel with the lowest natural cetane (and correspondingly highest specific gravity) is a virgin oil. The remaining neat biodiesels have relatively constant specific gravity, but widely varying natural cetane. Although not shown here, there was little variation in the other fuel properties for neat biodiesel.

The wide variation in natural cetane prompted two additional investigations. The first was to determine if the natural cetane number of biodiesel was an important component in the relationship between biodiesel concentration and emissions. This effort is described in more detail in Section IV.B.5. The second investigation was aimed at determining if natural cetane could be correlated with either plant or animal-based biodiesel categories. To do this, we plotted the distribution of natural cetane values separately for plant and animal-based neat biodiesel, and compared the two distributions. The results are shown in Figure II.E.1-2.

Figure II.E.1-2
Distribution of natural cetane for plant and animal-based neat biodiesel

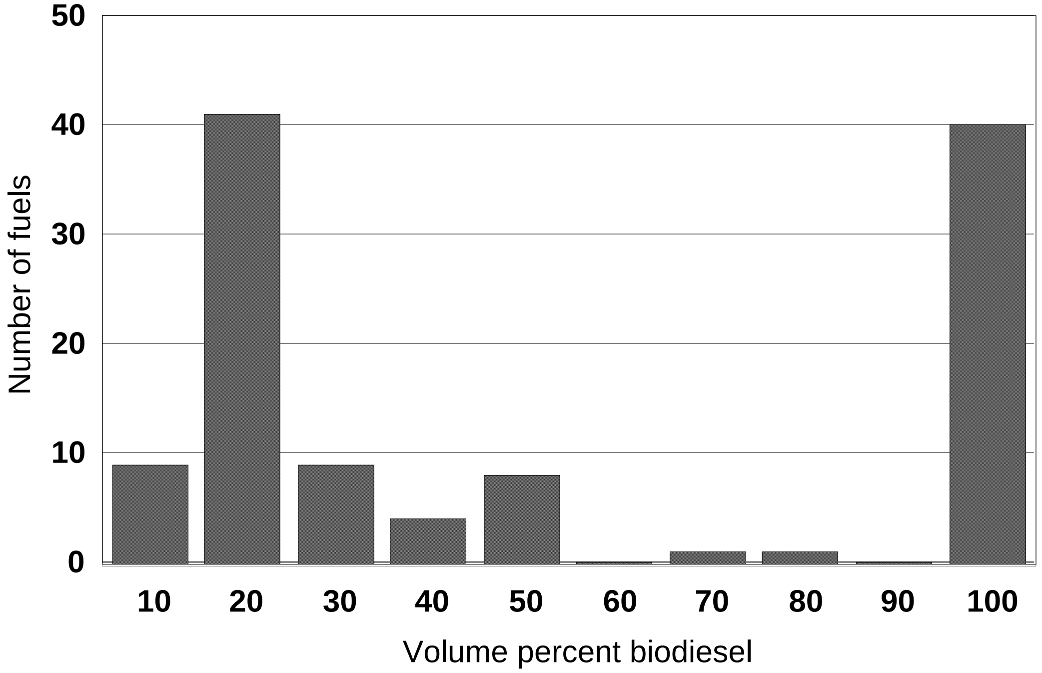


Based on a two-tailed t-test, the probability that plant and animal-based neat biodiesels in the database have different natural cetane values is 98%. Therefore, we have investigated whether the correlations between biodiesel concentration and emissions ought to be derived separately for animal and plant-based biodiesels.

There has also been some discussion in the literature about whether methyl esters differ from ethyl esters in terms of fuel properties and, ultimately, emission impacts. However, the studies included in our database often did not always specify whether the biodiesels in question were methyl or ethyl esters. A review of our database indicated that the cases where a clear distinction could be made were very small. Therefore, we determined that an investigation of the differences between methyl and ethyl esters was not possible in our analysis.

Biodiesel can be blended into conventional diesel fuel at any concentration. This fact is reflected in the distribution of biodiesel concentrations in the database, which is shown in Figure II.E.1-3.

Figure II.E.1-3
Distribution of biodiesel concentrations in database



Although 20 vol% biodiesel is the most common blend level among in-use biodiesel programs, the fact that biodiesel can be blended at any level suggests that the most useful analysis of biodiesel impacts on emissions would be to use biodiesel concentration as the independent variable in a traditional curve-fitting process. This approach would permit us to use all the available data in the analysis, and would provide a means for estimating the impact of biodiesel on emissions for any biodiesel concentration. Our curve-fitting approach is described in more detail in Section III.

2. Test cycles

When collecting data for input into our database, we excluded data that was collected on tests cycles that only contained a few nonstandard modes or which were otherwise deemed not representative of in-use operation. Table II.E.2-1 summarizes the number of observations in our database for each of the test cycles included in our analysis. The values in parentheses are the percent of total observations. "TRANSIENT" and "STEADY-STATE" tests cycles refer to all those cycles which were unique in some fashion, but which represented valuable data that could be used as validation sets.

Table II.E.2-1
Database observations by test cycle

Test cycle	HC	CO	NOx	PM	CO_2	BSFC
FTP composite	295 (36)[α]	295 (36)	295 (34)	294 (35)	175 (32)	167 (73)
FTP hot start	380 (46)	385 (47)	422 (49)	422 (51)	322 (59)	51 (22)
R49 13-mode	57 (7)	57 (7)	57 (7)	34 (4)	0	0
Nonroad 8 mode	14 (2)	14 (2)	14 (2)	8 (1)	0	0
"STEADY-STATE"	36 (4)	36 (4)	36 (4)	36 (4)	6 (1)	10 (4)
"TRANSIENT"	40 (5)	40 (5)	40 (5)	40 (5)	40 (7)	0
All cycles	822	827	864	834	543	228

[α] Values in parentheses are percent of total observations

3. Standards groups

As described in Section II.C above, categorizing the data in our database according to the emission certification standards that the engines were designed to meet might provide a convenient means for applying regression correlations to the in-use fleet. Table II.E.3-1 provides a summary of the number of engines and observations in our database.

Table II.E.3-1
Amount of data by engine standards groups for heavy-duty highway diesel engines

Standards group	Model years	HD highway engines	NOx observations
B	2002 - 2006	0	0
C	1998 - 2001	2	14 (2)[α]
D	1994 - 1997	10	152 (19)
E	1991 - 1993	16	394 (50)
F	1990	3	87 (11)
G	1988 - 1989	8	112 (14)
H	1984 - 1987	2	16 (2)
I	- 1983	2	10 (1)

[α] Values in parentheses are percent of total observations

Section III: How Was The Data Analyzed?

As mentioned previously, the goal of this study is to determine how diesel engine exhaust emissions are affected by the use of biodiesel. Although the most common biodiesel concentration is 20 volume percent, it can also be used at concentrations varying from 1 to 100 percent. As a result, we determined that a statistical regression correlation would be appropriate, offering a means for predicting the percent change in exhaust emissions as a function of the concentration of biodiesel in conventional diesel fuel.

This Section describes our statistical approach to estimating the effect of biodiesel on emissions of regulated pollutants. The results of applying these statistical approaches to the data in our database are described in the following two Sections for heavy-duty highway vehicles, and light-duty vehicles and nonroad engines. The impact of biodiesel on emissions of toxic pollutants is then presented separately in Section VI.

A. Overview of curve-fitting approach

1. Independent variables

We intended to correlate emissions as a function primarily of biodiesel concentration. Although it may have been ideal to include other fuel properties for biodiesel or the base fuel as independent variables, few of the studies which comprised our database included measurements of all relevant fuel properties. However, the category of "biodiesel" itself was considered to be a reasonable surrogate for the missing biodiesel fuel properties in light of the fact that biodiesel properties appeared to be largely constant across the studies in our database. In addition, we were able to investigate the impacts of base fuel properties in a general fashion based on a qualitative scale of "cleanliness," described more fully in Section III.C.2.e below.

While biodiesel fuel properties generally fell within a narrow range, natural cetane number was an exception; it varied significantly from one batch of biodiesel to another. Cetane number was measured for nearly every biodiesel and base fuel, providing a means for its inclusion in our analysis. However, since in the field the cetane number of a given batch of biodiesel or the base fuel to which biodiesel is added is not always known, correlations which include cetane number as an independent variable may not be the most user friendly. Therefore, although cetane number was not included in our final correlations, it was used in several other aspects of our analysis, such as:

- Establishing differences between animal and plant based biodiesel, as described in Section II.E.1
- Along with other fuel properties, categorizing base fuels as either "Clean" or "Average" emitting, as described more fully in Section III.C.2.e

- Determining whether EGR-equipped engines are likely to respond to biodiesel in a similar fashion to engines not equipped with EGR, as described in Section IV.B.5

One common technique in multivariable regression analysis is to standardize the independent variables. Standardization involves subtracting the mean from every observation, and then dividing the result by the standard deviation. It is useful for comparing the regression coefficients of the different independent variables to determine relative importance. However, in our approach we included only a single independent variable, the biodiesel concentration. Also, through preliminary regressions we determined that a squared biodiesel term was not necessary. As a result, we did not standardize the independent variable in our analysis.

2. Dependent variables

In reviewing the instances of repeat emissions data in our database (cases in which the same fuel was tested on the same engine multiple times), it appeared that the variability in emissions measurements was a function of the mean emissions measurements. In other words, lower emission levels tended to exhibit smaller variability than higher emission levels. As a result, we determined that the use of a log transform for emissions would be appropriate, since its use tends to make the dependent variable's data in a regression equation homoscedastic.

The use of a log transform has another advantage. Our analysis was intended to produce correlations that predict the percent change in emissions resulting from the use of a given concentration of biodiesel. The analysis was not intended to permit the estimation of absolute emission levels (g/mile or g/bhp-hr). Given the relative nature of the correlations, the intercept terms produced during the regressions can be eliminated. This result is shown mathematically below.

The regression equation can be expressed as:

$$\log(\text{Emissions}) = a \times (\text{vol\% biodiesel}) + b \qquad (1)$$

where a and b are determined through the statistical curve-fitting process. In practice, absolute emissions would be estimated from:

$$\begin{aligned} \text{Emissions} &= \exp[a \times (\text{vol\% biodiesel}) + b] && (2)\\ &= \exp[a \times (\text{vol\% biodiesel})] \times \exp(b) && (3) \end{aligned}$$

The percent change in emissions due to the use of biodiesel is calculated generally from:

$$\text{\% change in emissions} = \frac{(\text{Emissions})_{\text{with biodiesel}} - (\text{Emissions})_{\text{without biodiesel}}}{\text{Emissions}_{\text{without biodiesel}}} \times 100 \qquad (4)$$

Equation (3) can be combined with (4) to produce:

$$\%\text{ change in emissions} = \frac{\{\exp[a \times (\text{vol\% biodiesel})] \times \exp(b) - \exp[a \times 0] \times \exp(b)\}}{\exp[a \times 0] \times \exp(b)} \times 100 \tag{5}$$

$$\%\text{ change in emissions} = \frac{\{\exp[a \times (\text{vol\% biodiesel})] - 1\} \times \exp(b)}{\exp(b)} \times 100 \tag{6}$$

$$\%\text{ change in emissions} = \{\exp[a \times (\text{vol\% biodiesel})] - 1\} \times 100 \tag{7}$$

Equation (7) does not contain the constant b, and can be used to predict the percent change in emissions for a given concentration of biodiesel.

3. Curve fitting approach

Rather than using a least-squares type of regression, we opted to use a maximum likelihood approach to curve fitting. In SAS, this approach is employed with the procedure proc_mix. This procedure is less prone than least-squares to being influenced by large numbers of repeat measurements. It can also treat some variables as fixed effects and others as random effects. For instance, the primary independent variable that we intended to include in the correlations, percent biodiesel, was represented as a fixed effect. Other variables, such as engines and base fuels, were based on data that is a sampling from a wider population. As such, they are best represented as random effects. Our curve-fitting effort, therefore, included engine intercepts, engine × percent biodiesel, and engine × base fuel terms as random effects.

As the analysis progressed, we used several criteria for determining when candidate fixed terms should be included in the correlations. The first was a screening tool that ensured that a minimum amount of data was available for any adjustment term considered for inclusion in the correlations. This screening tool is described in more detail in Section III.C.1 below. We also used a significance criterion of $p = 0.05$ for all runs. We did not make use of Mallow's Cp criterion to balance over-fitting and under-fitting, since this criterion cannot be calculated in a mixed effects correlation. However, the number of adjustment terms added to our final correlations was quite small and would be unlikely to cause significant overfitting.

B. Treatment of different types of diesel equipment

The database includes emissions data on a variety of diesel equipment types. These include highway and nonroad, light-duty and heavy-duty, and both engines and vehicles. There was no biodiesel data available for stationary source engines. The distribution of NOx observations between these various categories is given in Table III.B-1.

Table III.B-1
Distribution of NOx observations by diesel equipment category

	NOx observations	Percent of observations
Heavy-duty highway engines	658	80
Heavy-duty highway vehicles	143	17
Heavy-duty nonroad engines	14	2
Light-duty highway engines	6	1

We decided to focus our initial curve-fitting efforts on the data from heavy-duty highway engines, since this category contained the most data. We did not include the heavy-duty highway vehicle data in this initial analysis, since we did not want to presume that biodiesel effects on emissions would be the same for engines and vehicles, given differences in the associated test cycles. Instead, once we had completed our investigations of heavy-duty highway engines, we compared the resulting correlations to the data from the remaining three categories of diesel equipment to determine how well the correlations represented these three other categories. Later sections of this report describe these comparisons.

We also intended to examine the possibility that engine technology might be an important factor in correlating biodiesel use with emissions. However, the engine descriptions in the studies that comprise our database were often incomplete, making it difficult to assign each engine to any but the broadest of technology groupings. This was not deemed a significant detriment to our analysis for two reasons:

1. A previous analysis of the impacts of diesel fuel properties on emissions[3] concluded that adjustment terms representing different engine technologies were rarely necessary.
2. Comments on our previous analysis had suggested that grouping engines by model year might be more appropriate than grouping them by technology type.

As a result, we grouped all engines into one of seven "standards groups," based on their model year which was available for nearly every engine in our database. These standards groups are described in Section II.C, and the distribution of NOx data among these groups is described in Section II.E.3.

There was one important group of engine technologies which was notably missing from our database: engines equipped with exhaust gas recirculation (EGR), designed to meet the 2004 heavy-duty engine certification standards. Because these engines will comprise a larger and larger fraction of the in-use fleet in the coming years, their absence from our database raises the question of how and to what degree our correlations should apply to the in-use fleet. However, we note that in a previous analysis it was primarily cetane effects that were different for EGR-equipped engines as compared to non-EGR engines. Specifically, EGR-equipped engines appeared to exhibit no NOx response to changes in cetane number, whereas non-EGR engines exhibited reductions in NOx when cetane number increased. Thus, although no EGR-equipped engines were in our biodiesel database, we designed an analytical approach that provided insight into how NOx emissions from EGR-equipped engines might respond to the use of biodiesel. This analysis is described in Section III.C.2.f.

C. Inclusion of second-order and adjustment terms

We investigated the need for additional terms in our correlations in order to increase their explanatory power. These additional terms included a squared term for biodiesel concentration and adjustment terms for test cycle, biodiesel source, engine standards groups, and the conventional diesel fuel to which biodiesel was added. We also investigated whether cetane number was an important element in the correlation of biodiesel concentration with emissions. Our approach to investigating the need for these additional terms is described in this section.

In response to previous work correlating diesel fuel properties with emissions, stakeholders suggested that we develop criteria that establish the minimum amount of data that would be necessary before a given adjustment term should even be considered for inclusion in statistical correlations. Such criteria would provide some insurance against statistically significant adjustment terms entering the correlations as a result of a limited amount of data that just happens to exhibit a spurious emissions effect. Therefore, Section III.C.1 describes the development and application of these minimum data criteria. Section III.C.2 will describe the various specific adjustment terms that we investigated.

1. Minimum data criteria

The statistical significance of any potential adjustment term that we could add to our correlations depends in part on the number of observations that represent that adjustment term. For a small sample, the potential exists for the mean effect to significantly differ from the true population mean. For instance, there is a 13% probability that every observation in a sample of four observations would fall on one side of a normal distribution curve. For a sample of eight observations, the probability of this occurring drops to less than 1%. Similarly, a data set that is too small may lead to the conclusion that an adjustment term in our correlations is statistically significant when in fact a larger sample would prove otherwise. Thus it seemed prudent to establish some minimum number of observations below which we would not consider including a potential adjustment term in our correlations.

We investigated criteria for the minimum amount of data that would be needed to estimate a population mean within some confidence interval. The formula[4] for this calculation is:

$$n = \left(\frac{z_{\alpha/2} \times \sigma}{E} \right)^2 \qquad (8)$$

where

n = Minimum number of observations
$z_{\alpha/2}$ = z-value at a confidence level of 1-α
σ = Population standard deviation (in g/bhp-hr)
E = One-half of the confidence interval (in g/bhp-hr)

For a confidence level of 90%, the value of $z_{\alpha/2}$ is 1.645. We estimated the population standard deviation from the emission measurements in a database developed for earlier work (see summary in our July 2001 Staff Discussion Document). These standard deviations were calculated by first subtracting the mean from all repeat measurements (occasions in which the same fuel was tested on the same engine multiple times), resulting in 296 differences for each pollutant. We then calculated the standard deviation for these differences. The results are shown in Table III.C.1-2 below. We did not simply calculate the standard deviation from the original emission measurements, because different engines were designed to meet different standards and so would have produced an overall standard deviation not representative of repeat measurements of a single fuel on a single engine.

In order to estimate the minimum number of observations, we also need to estimate the confidence interval in g/bhp-hr defining the maximum error we are willing to accept. In equation (8), E represents one-half of the 90% confidence interval, measured in g/bhp-hr, within which we would consider an adjustment term to be no different than the overall biodiesel concentration term, and outside of which inclusion of an adjustment term in our correlations might be warranted. In order to estimate E, we needed to make some preliminary estimates of the impact of biodiesel on emissions in percent change, choose an interval around these percent change estimates to represent the maximum error we might find acceptable, and then convert this interval into g/bhp-hr. This calculation for E was based on the following equation:

$$E(\text{g/bhp-hr}) = |\%\ \text{change}| \times E(\text{percent}) \times \text{Mean emissions (g/bhp-hr)} \qquad (9)$$

where

E(g/bhp-hr) = One-half the width of the confidence interval, used in equation (8)
% change = Predicted % change in emissions due to the addition of biodiesel
E(percent) = One-half the width of the confidence interval as a fraction of the effect of biodiesel on emissions
Mean emissions = Average g/bhp-hr for conventional diesel fuel

We chose the % change values to represent a blend of 20% biodiesel, currently the most common biodiesel concentration. Our preliminary correlations predicted that the percent change in emissions for 20% biodiesel for HC, CO, NOx, and PM were -20, -10, +2, and -10, respectively. We arbitrarily chose an interval of ± 25% around each of these percent change values to define the confidence interval E(percent). Thus, for instance, the confidence interval for HC would be -15% to -25%, and this range of biodiesel effects on HC was converted into a value of E in units of g/bhp-hr using equation (9). The mean emissions values for each pollutant were drawn from the same database used to estimate the population standard deviations. Table III.C.1-1 summarizes the calculation of E(g/bhp-hr) for all four pollutants.

Table III.C.1-1
Calculation of confidence interval for minimum data criteria

	HC	CO	NOx	PM
% change for B20	-20	-10	+2	-10
E (percent)	± 25 %	± 25 %	± 25 %	± 25 %
Mean emissions (g/bhp-hr)	0.25	0.96	4.87	0.11
E (g/bhp-hr)	±0.012	±0.024	±0.024	±0.0029

The calculation of the minimum number of observations then follows equation (8). The results are given in Table III.C.1-2.

Table III.C.1-2
Minimum number of observations

	HC	CO	NOx	PM
Population standard deviation (g/bhp-hr)	0.024	0.052	0.072	0.0058
E (g/bhp-hr)	0.012	0.024	0.024	0.0029
n	10	12	23	11

The values for n in Table III.C.1-2 represent reasonable lower limits for the number of observations that would be necessary in order to have some confidence that statistically significant adjustment terms are legitimate. However, we note that the estimated value of n for NOx is considerably higher than that for the other pollutants, owing primarily to the smaller impact that biodiesel has on NOx emissions. Rather than have separate minimum data criteria for each pollutant, and recognizing that alternative values of n could reasonably be estimated with different inputs to equation (8), we have decided to use a single value of 20 observations as our minimum data criterion.

We made two modifications to the use of a minimum data criterion of 20 in our curve-fitting effort. First, because our analysis was intended to permit estimation of the percent change in emissions resulting from the use of biodiesel, a minimum of two observations are required to establish each point estimate: the base fuel, and the biodiesel blend. Therefore, we applied our minimum data criteria to point estimates, i.e. pairs of observations consisting of a base fuel and a biodiesel blend. In practice, this meant counting only biodiesel blends when making a determination as to whether the minimum data criterion had been met.

Second, we determined that any subset of data being considered as the basis for an adjustment term should contain at least two engines. This additional criterion reduced the chances that an engine with unique responses to biodiesel would by itself form the basis of an adjustment term in any of our correlations.

2. Curve-fitting approach for specific terms

This section describes the various adjustment terms that we considered adding to our correlations, including the application of our minimum data criteria. The analytical approaches taken are described here, while results of the analyses are described in Section IV. All of these analyses were done only for heavy-duty highway engines, after which we made comparisons of the resulting correlations to data for other types of diesel equipment.

The evaluation of every type of potential adjustment terms (test cycle effects, biodiesel source effects, etc.) was initially done independently from all other types of potential adjustment terms. In each case, the correlations were generated in three steps:

Step 1: Generate a correlation to identify outliers
Step 2: After dropping outliers, generate a correlation to identify statistical significance of terms
Step 3: After dropping non-significant adjustment terms, generate a final correlation

Once all the important adjustment terms had been identified through this process, a single correlation incorporating them all was developed. These 'composite correlations' are described in Section IV.B.6.

a. Squared biodiesel term

There were a number of studies in our database that tested three or more different biodiesel concentrations using a single base diesel fuel. We therefore felt it appropriate to investigate whether a squared biodiesel concentration term should be added to our correlations. For all pollutants, the squared biodiesel concentration term was not significant. Thus for all subsequent analysis, only a linear biodiesel concentration term was included.

b. Test cycle effects

The database contained emission measurements collected on a variety of test cycles, as shown in Table II.E.2-1. For our initial analysis of heavy-duty highway engines, we set aside all data collected on generic transient or generic steady-state cycles. This left only composite FTP, hot-start FTP, and R49 data. The distribution of biodiesel NOx observations for this data subset is shown in Table III.C.2.b-1.

Table III.C.2.b-1
Biodiesel NOx observations by cycle for heavy-duty highway engines

FTP composite	135
FTP hot start	175
R49 13-mode	18

Although the number of R49 observations did not strictly meet our minimum data criterion of 20 observations, previous analyses suggested that test cycle may have an effect on emissions from heavy-duty diesel engines, particularly for PM and CO. Since the R49 data was collected on three separate engines, and comes close to our minimum data criterion, in this case we decided to proceed with investigating the need for adjustment terms representing all three test cycles. Thus we did not automatically exclude the steady-state R49 data from the PM and CO analyses, but instead investigated whether or not it was appropriate to include this steady-state data in our analysis of biodiesel effects on PM and CO emissions.

In addition to biodiesel concentration as a fixed effect in the SAS procedure proc_mix, we introduced terms representing each of the three test procedures and interactions between them and biodiesel concentration. The list of fixed effects included in this analysis are shown in Table III.C.2.b-2.

Table III.C.2.b-2
Fixed terms used to investigate test cycle effects

- Overall intercept
- R49 intercept
- UDDS intercept
- UDDSH intercept
- Percent biodiesel
- Percent biodiesel × R49
- Percent biodiesel × UDDS
- Percent biodiesel × UDDSH

Because the number of interactive terms for which coefficients can be estimated in the fixed portion of proc_mix is limited to n-1, where n is the number of test cycle groups, no coefficient could be estimated for one of the three interactive terms. In this case, no estimate was made for the percent biodiesel × UDDSH interactive term. As a result, the overall percent biodiesel term defaulted to representing the UDDSH cycle. The results of our investigation of test cycle effects is described in Section IV.B.1.

c. Biodiesel source effects

Biodiesel can be produced from a wide variety of feedstocks. The studies that comprise our database included only a portion of the many feedstocks possible, though they do represent the most common feedstocks. The biodiesel feedstocks in our database are listed in Table III.C.2.c-1.

Table III.C.2.c-1
Biodiesel feedstocks in database

Feedstock	Number of biodiesel observations
Soybeans	232
Rapeseeds	41
Canola oil	3
Grease	23
Tallow	9
Lard	3

From Table III.C.2.c-1 we see that at least three feedstock categories do not contain sufficient data for an analysis of feedstock impacts on the correlation between biodiesel concentration and emissions. We therefore investigated ways in which the feedstocks in Table III.C.2.c-1 could be combined into larger groups. We know from a review of pure biodiesel cetane numbers that plant-based biodiesel (soybean, rapeseed, and canola) are distinct from animal-based biodiesel (grease, tallow, and lard). We also know that canola oil is derived from a rape plant offbreed, and so might be appropriately combined with rapeseed-based biodiesel. As a result, we created three groups which are listed in Table III.C.2.c-2. Every biodiesel blend in our database was placed into one of the three source groups shown in this table.

Table III.C.2.c-2
Biodiesel source groups

Feedstock	Number of biodiesel observations
Soybeans	232
Rapeseeds/canola	44
All animal	35

In addition to biodiesel concentration as a fixed effect in the SAS procedure proc_mix, we introduced terms representing each of the three biodiesel source groups and interactions between them and biodiesel concentration. The list of fixed effects included in this analysis are shown in Table III.C.2.c-3.

Table III.C.2.c-3
Fixed terms used to investigate biodiesel source effects

Overall intercept
Soybean intercept
Rape intercept
Animal intercept
Percent biodiesel
Percent biodiesel × soybean
Percent biodiesel × rape
percent biodiesel × animal

Because the number of interactive terms for which coefficients can be estimated in the fixed portion of proc_mix is limited to n-1, where n is the number of biodiesel source groups, no coefficient could be estimated for one of the three interactive terms. In this case, no estimate was made for the percent biodiesel × soybean interactive term. As a result, the overall percent biodiesel term defaulted to representing soybean-based biodiesel. The results of our investigation of our investigation of biodiesel source groups is described in Section IV.B.2.

d. Effects of engine standards groups

As described in Sections II.C and III.B, we grouped all heavy-duty highway engines into one of seven "standards groups," based on their model year which was available for nearly every engine in our database. These standard groups were used as a surrogate for engine technology for which data was largely missing. We then investigated whether the biodiesel effects on emissions were significantly different between each of these standards groups. The number of NOx observations for biodiesel blends is given in Table III.C.2.d-1.

Table III.C.2.d-1
Number of biodiesel NOx observations for heavy-duty highway diesel engines

Standards group	Model years	NOx observations
C	1998 - 2001	0
D	1994 - 1997	113
E	1991 - 1993	149
F	1990	15
G	1988 - 1989	44
H	1984 - 1987	2
I	- 1983	4

This table shows that three of the standards groups, F, H, and I, do not meet our minimum data criteria for the investigation of subgroup effects of biodiesel on emissions. We therefore only investigated adjustment terms for engine standard groups D, E, and G.

In addition to biodiesel concentration as a fixed effect in the SAS procedure proc_mix, we introduced terms representing each of the three engine standards groups and interactions between them and biodiesel concentration. The list of fixed effects included in this analysis are shown in Table III.C.2.d-2. The results of our investigation of engine standard group effects is described in Section IV.B.3.

Table III.C.2.d-2
Fixed terms used to investigate engine standards group effects

Overall intercept
Group D intercept
Group E intercept
Group G intercept
Percent biodiesel
Percent biodiesel × Group D
Percent biodiesel × Group E
Percent biodiesel × Group G

e. Base fuel effects

When biodiesel is added to a base fuel, we initially assumed that any impacts on emissions would be strictly a function of the biodiesel concentration. That is, the impact that biodiesel has on emissions at a given concentration would be the same regardless of the properties of the base fuel to which the biodiesel was added. However, it seemed prudent to test this assumption.

Ideally, the base fuel would be taken into account in our curve-fitting process by adding terms to the correlations representing the properties of the base fuel, in addition to the biodiesel concentration term. The resulting correlations would then allow one to predict the impact that biodiesel has on emissions both as a function of the biodiesel concentration and the properties of the base fuel, if the terms for those base fuel properties were statistically significant. Unfortunately, many of the studies which comprise our database did not include a full complement of fuel property measurements for the base fuel. Of the 39 base fuels, only 21 had measured values for all the primary fuel properties of interest (cetane, sulfur, aromatics, T10, T50, T90, density). These fuel properties were identified in our July 2001 Staff Discussion Document as being important measures of fuel property effects on emissions of regulated pollutants.

In order to maximize the useable data when investigating base fuel effects, we opted to divide all the base fuels into one of three groups: clean, average, and dirty. The assignments were made using a combination of available measured fuel properties and the description of the base fuel from the original study. Using a distribution of fuel properties from an in-use diesel fuel survey, we determined that none of the base fuels in our database should be assigned a label of "dirty." Many base fuels were designed to meet certification fuel standards, and these were assigned to the "average" group. The complete assignments are given in Appendix C. Generally speaking, the "clean" fuels had lower aromatics, higher cetane, lower density, and lower distillation points than "average" fuels.

Based on the distribution of fuel properties for these two base fuel groups, we here propose a set of conditions that might permit one to distinguish a "clean" base fuel from an "average" base fuel when using the correlations presented in Section IV. These conditions are given in Table III.C.2.e-1.

Table III.C.2.e-1
Base fuel emission group proposed definitions

A. All base fuels to which biodiesel is added are assigned to the "average" emission category for the purposes of estimating emission benefits of biodiesel using the correlations in this report, unless
B1. The base fuel in question meets the requirements for highway diesel fuel sold in California or alternative requirements that are substantially similar to those in California, or
B2. The fuel in question meets all of the following conditions: 1. Total cetane number is greater than 52 2. Total aromatics content is less than 25 vol% 3. Specific gravity is less than 0.84 For fuels meeting conditions B1 or B2, the base fuel should be assigned to the "clean" category.

In addition to biodiesel concentration as a fixed effect in the SAS procedure proc_mix, we introduced terms representing each of the two base fuel emission groups and interactions between them and biodiesel concentration. The list of fixed effects included in this analysis are shown in Table III.C.2.e-2.

Table III.C.2.e-2
Fixed terms used to investigate base fuel effects

Overall intercept
Average intercept
Clean intercept
Percent biodiesel
Percent biodiesel × average
Percent biodiesel × clean

Because the number of terms for which coefficients can be estimated in the fixed portion of proc_mix is limited to n-1, where n is the number of base fuel emission groups, no coefficient could be estimated for one of the two interactive terms. In this case, no estimate was made for the percent biodiesel × clean interactive term. As a result, the overall percent biodiesel term defaulted to representing 'clean' base fuels. The results of our investigation of base fuel effects in given in Section IV.B.4.

f. Cetane effects

As shown in Table II.E.3-1, our database does not contain any data on model year 2002+ heavy-duty highway engines. Heavy-duty highway engines produced to meet the 2004 standards (and which, under a consent decree, may be produced as early as October of 2002) are expected to make substantial use of exhaust gas recirculation (EGR). In a previous analysis[5] we concluded that EGR-equipped engines respond differently to the cetane number of diesel fuels, in terms of NOx emissions, than engines without EGR. In short, cetane appears to have a significant role in NOx emission effects for engines that do not have EGR, but a negligible role in NOx emission effects for EGR-equipped engines.

As shown in Figure II.E.1-1, the cetane number of biodiesel varies widely. On average, however, the natural cetane number of biodiesel is higher than that for conventional diesel fuel (see Table II.E.1-1). If the cetane number of biodiesel plays a role in the impact that biodiesel has on NOx emissions, it is possible that EGR-equipped engines may exhibit a different biodiesel/NOx relationship than engines not equipped with EGR. In order to test this hypothesis, we conducted an analysis to determine the importance of cetane in the correlation of biodiesel concentration with NOx emissions. This analysis was not done in order to add a cetane term to the NOx correlation, but rather to determine if it was legitimate to extrapolate the effects of the NOx correlation to EGR-equipped engines.

Through previous work, we had developed a draft[c] correlation between cetane number and NOx emissions. This correlation is shown below:

$$\text{NOx} = \exp(-0.004512 \times \text{CET\_DIF} + 0.0001458 \times \text{CET\_DIF}^2 + 1.5497034)$$

The correlation was based on NOx differences between fuels containing a cetane improver additive and their associated unadditized base fuel. The independent variable CET_DIF is the difference between the cetane number of the additized fuel and the cetane number of the base fuel. As a result, this correlation avoids the collinear effects that natural cetane often has with aromatics and specific gravity.

In order to determine the degree to which cetane plays a role in the biodiesel/NOx relationship, we used the above equation to separate cetane effects in the biodiesel database from the measured NOx emissions. To do this, we first converted the above equation into an equivalent form that could be applied to cetane number (as opposed to cetane difference). The result is shown below:

---

[c] This correlation was generated in preliminary analyses and would not be appropriate for use in other contexts. EPA has released a draft technical report that contains our updated analysis on additized cetane effects on NOx emissions: "The effect of cetane number increase due to additives on NOx emissions from heavy-duty highway engines," EPA Draft Technical Report EPA420-S-02-012, June 2002.

$$NOx = \exp(-0.004512 \times (CETNUM - 45) + 0.0001458 \times (CETNUM - 45)^2 + 1.5497034)$$

where CETNUM is the total cetane number of any fuel. We then used this equation to predict the NOx emissions for every fuel in our biodiesel database, and calculated the difference between the predicted NOx and measured NOx for every observation. These "pseudo-residuals" should therefore be independent of any cetane effects, and could be used as dependent variables in a mixed model analysis of biodiesel concentration. Depending on the magnitude of the biodiesel effect for this correlation and its associated statistical significance, we could determine the degree to which cetane plays a role in the relationship between biodiesel concentration and NOx emissions. The results of this analysis are described in Section IV.B.5.

Section IV: Biodiesel Effects on Heavy-Duty Highway Engines

Our primary analyses were carried out only on heavy-duty highway engine data, since this data comprised the majority of the database. In this context we evaluated the need for various adjustment terms and produced a single correlation for each of the pollutants NOx, PM, CO, and HC which included all those effects that were deemed significant. We also investigated fuel economy impacts of biodiesel and the impact that biodiesel use would have on emissions of the greenhouse gas carbon dioxide. The investigation of toxics impacts of biodiesel use followed an alternative analytical approach, and so is covered separately in Section VII.

This section describes the results of our analysis for heavy-duty highway engines, and shows how well on-highway heavy-duty vehicle emissions data represents engine impacts of biodiesel use. In Section V we compare the correlations presented in this Section to data collected on light-duty vehicles and nonroad engines.

A. Basic correlations

Before investigating the degree to which biodiesel effects differ due to engine technology, type of biodiesel, type of base fuel, or test cycle, we first carried out more fundamental analyses of biodiesel effects on regulated pollutants, fuel economy, and carbon dioxide. This Section summarizes our analyses of these basic correlations, while Section IV.B provides a description of the various adjustments we investigated.

1. Regulated pollutants

The first step in our analysis involved development of correlations between biodiesel concentration and the percent change in emissions, without any adjustment terms[d], for heavy-duty highway engines. All correlations were of the form:

$$\text{\% change in emissions} = \{\exp[a \times (\text{vol\% biodiesel})] - 1\} \times 100\%$$

Table IV.A-1 gives the values for the coefficient "a" for each of the four pollutants, all of which were statistically significant at the $p = 0.05$ level. Figure IV.A.1-1 presents these basic correlations graphically.

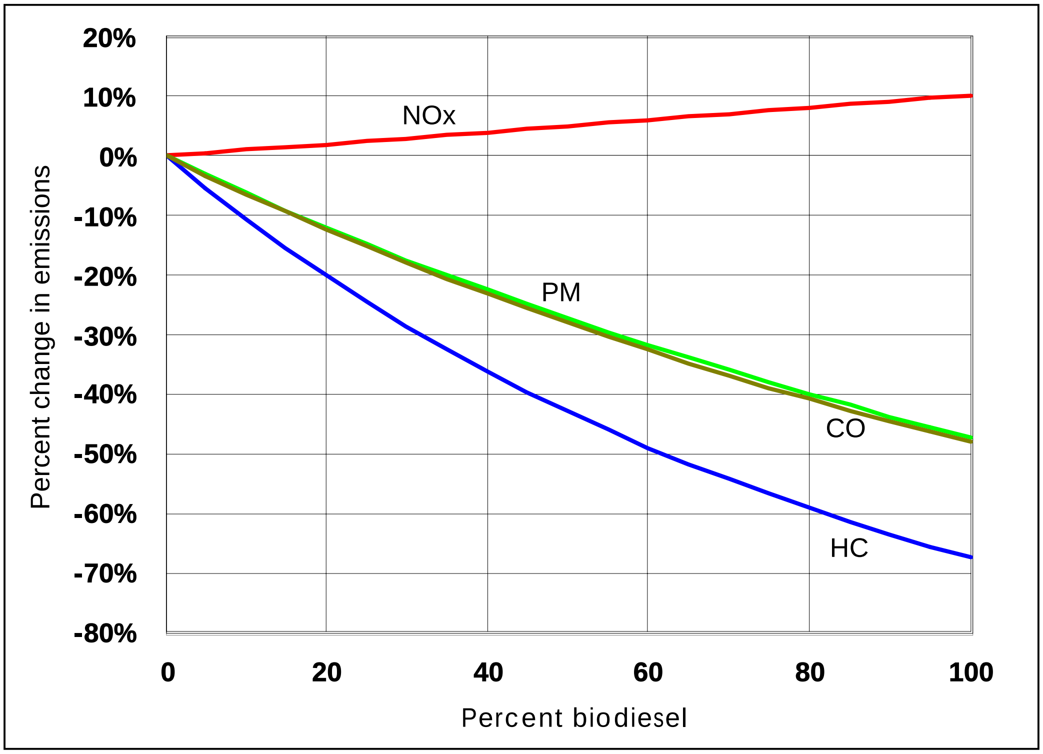
---

[d] Subsequent analyses indicated that R49 data should not be included in curve-fitting for PM or CO. See Section IV.B.1. Therefore, R49 data was removed from the basic correlations for PM and CO.

Table IV.A.1-1
Coefficients for basic emission correlations

	Coefficient "a"
NOx	0.0009794
PM	-0.006384
HC	-0.011195
CO	-0.006561

Figure IV.A.1-1
Basic emission correlations



Because we used maximum likelihood curve-fitting instead of least-squares regressions to develop our correlations, the coefficient of determination (r^2) is not applicable and was not provided by the SAS procedure proc_mix. As an alternative, we plotted the actual data on the same graph as our correlations to provide a visual assessment of goodness of fit. To do this, we needed to calculate the percent change values associated with every observation in our database. Because the number of repeat tests on base fuel (0% biodiesel) were not always paired with an equivalent number of tests on biodiesel, this process required that we first calculate the average of repeat emission measurements made on every base fuel (0% biodiesel) for every engine/test cycle combination. Percent change in emission values for every biodiesel blend were then calculated using this average base fuel. The comparisons are shown in Figures IV.A.1-2, IV.A.1-3, IV.A.1-4, IV.A.1-5 for NOx, PM, HC, and CO, respectively.

Figure IV.A.1-2
Comparison of data to basic NOx correlation

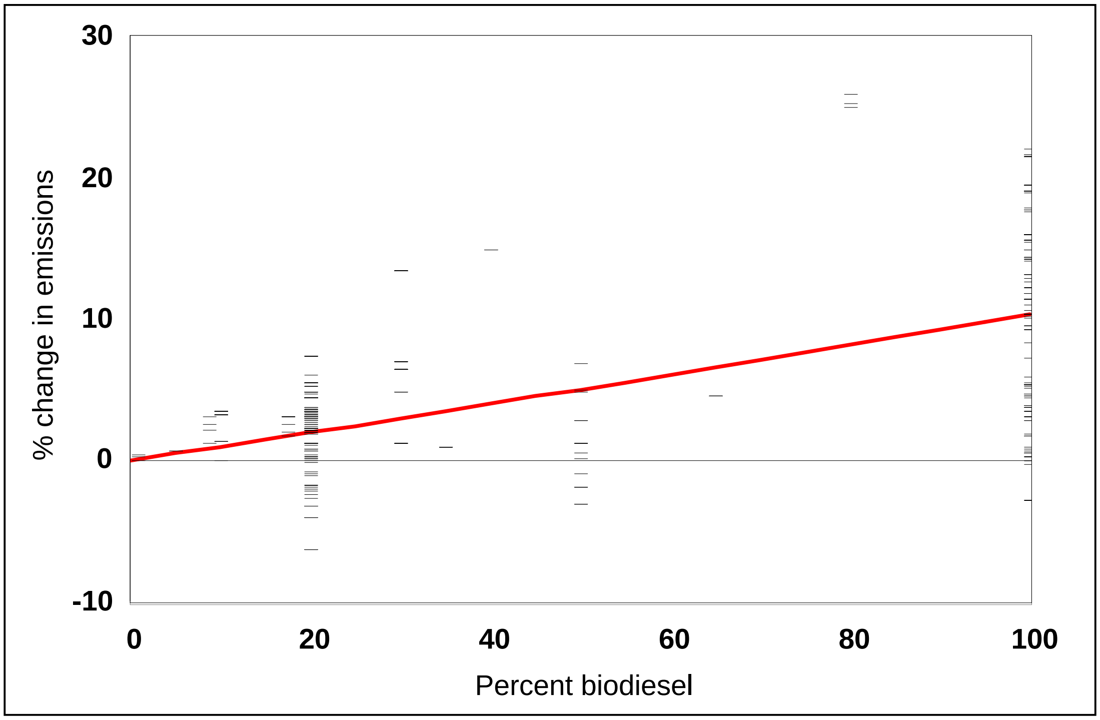


Figure IV.A.1-3
Comparison of data to basic PM correlation

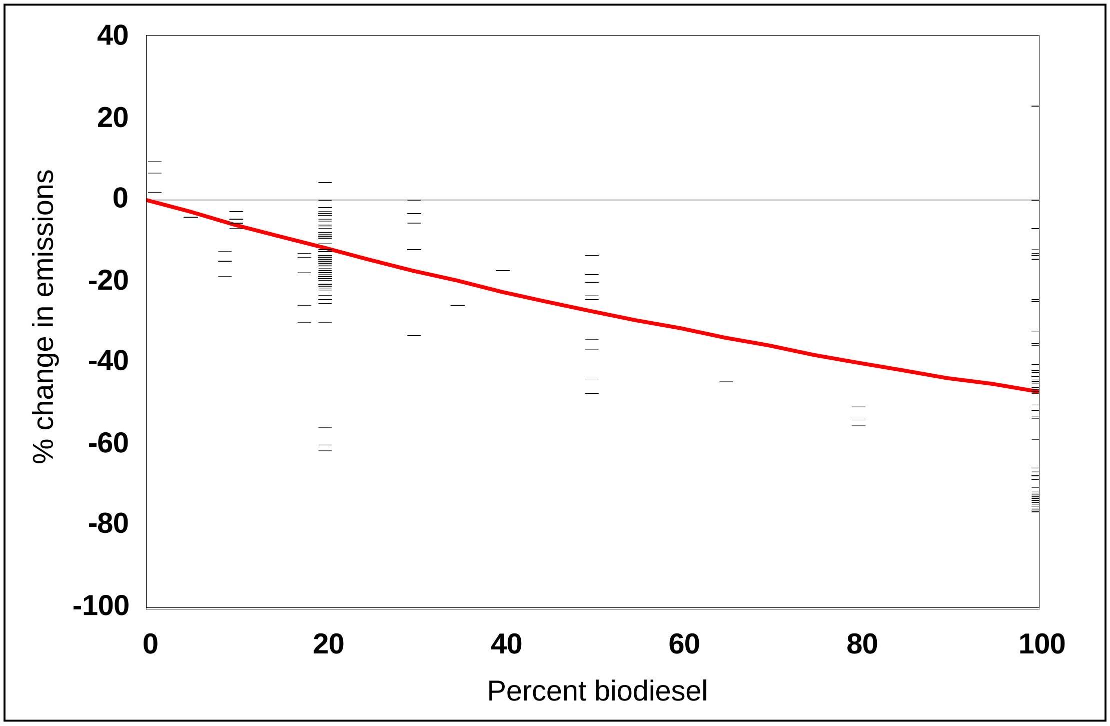


Figure IV.A.1-4
Comparison of data to basic HC correlation

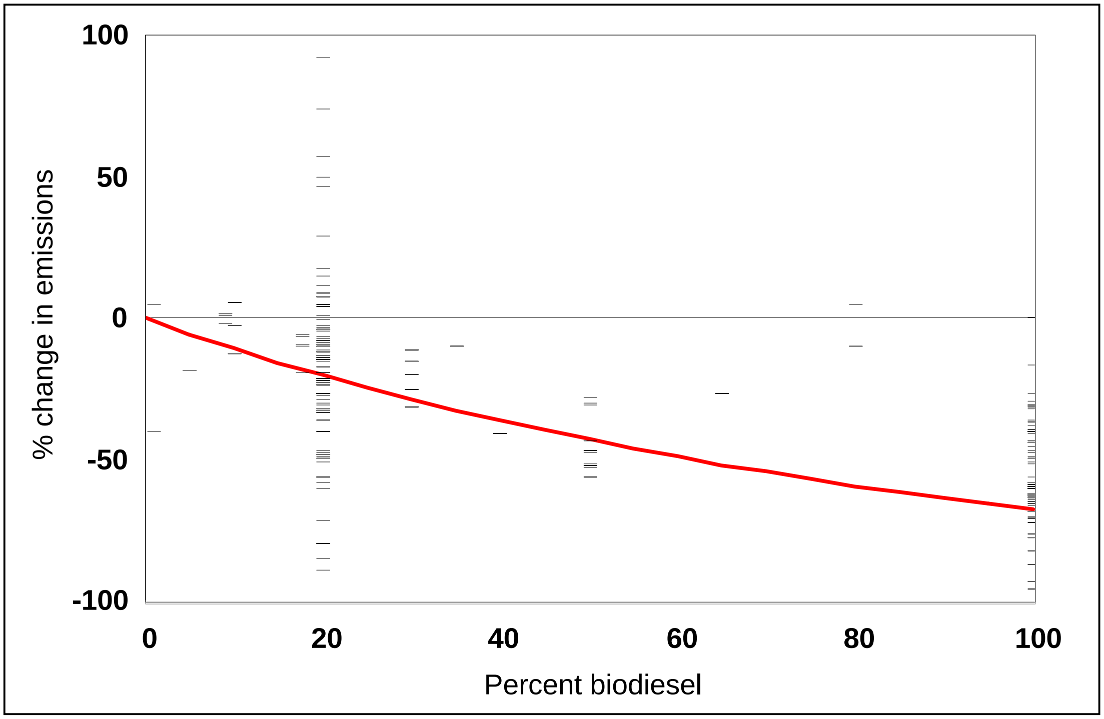
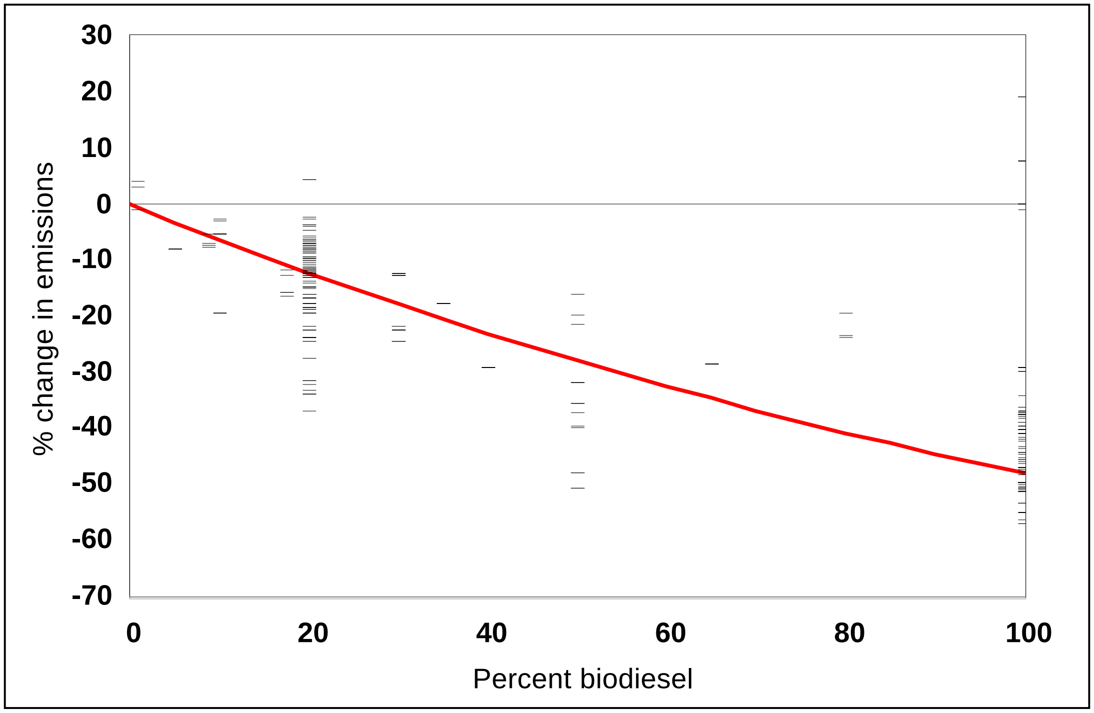


Figure IV.A.1-5
Comparison of data to basic CO correlation



Emissions of all pollutants except NOx appear to decrease when biodiesel is used. The fact that NOx emissions increase with increasing biodiesel concentration could be a detriment in areas that are out of attainment for ozone. Thus it seemed appropriate to consider the limited conditions under which NOx might actually decrease with increasing biodiesel concentration. As can be seen in Figure IV.A.1-2, a number of observations in our database actually exhibit a decrease in NOx when biodiesel is added to conventional diesel fuel. Some such results are possible simply due to test measurement variability. For instance, using the values for mean NOx emissions and the standard deviation from Tables III.C.1-1 and III.C.1-2, we were able to estimate that repeated measurements of a 20% biodiesel blend could actually show a decrease in NOx emissions 10 percent of the time even though the mean effect might indicate a 2% increase in NOx. Still, there might be other elements that might cause NOx emissions to decrease with increasing biodiesel concentration. We therefore took a closer look at these observations to determine if there was anything unique about them. Table IV.A.1-2 summarizes our review of these observations.

Table IV.A.1-2
Review of observations for which NOx emissions decrease with increasing biodiesel

Conclusion	Implication
All observations were tested using soy-based biodiesel.	None. 75% of the biodiesel observations in the database are soy-based.
Base fuels have an average cetane number of 49.	The average is influenced by a single Fischer-Tropsch base fuel. The remaining six base fuels have an average cetane number of 45, equivalent to average in-use conventional diesel fuel.
Only engine standard groups D and E are represented.	None. 80% of the observations in the database were collected on standard groups D and E
Of the eight engines, various manufacturers, engine sizes, and injection control systems are represented.	Nothing unique about the engines.

Based on this review, there is no consistent set of characteristics for these eight engines, the type of biodiesel used, and the base fuels to which biodiesel was added that would explain why NOx emissions decreased when biodiesel replaced conventional diesel. However, it is notable that these engines all exhibited more beneficial emission effects due to biodiesel use for all pollutants in comparison to other engines in our database. This result implies that these eight engines respond to biodiesel use in some fundamentally different, albeit subtle, way than other engines even though the characteristic(s) that produce this result are not readily apparent. This issue may warrant further investigation.

2. Fuel economy impacts of biodiesel use

Biodiesel use will generally reduce the number of miles that a vehicle can be driven on a gallon of fuel due to its lower energy content in comparison to conventional diesel fuel. We used two different approaches to quantify this reduction in fuel economy: differences in fuel energy content between pure biodiesel and conventional diesel, and correlations between biodiesel concentration and brake-specific fuel consumption.

a. Via fuel energy content

Our database contained 19 pure biodiesel fuels for which energy content was measured. Generally energy content was reported in terms of net energy per unit mass in the studies. We converted these units into Btu/gal using each fuel's measured specific gravity to best represent the impact on fuel economy. We also used only the net (lower) heating value of the fuels, since this is a better approximation of the available work that can be extracted from the fuel than the gross (higher) heating value. The average values are shown in Table IV.A.2.a-1.

Table IV.A.2.a-1
Average energy content of 100% biodiesel

	Average net Btu/gal
All biodiesels	118,296
Animal-based	115,720
Rapeseed/canola-based	119,208
Soybean-based	119,224
Rapeseed or soybean-based	119,216

A two-tailed t-test of the biodiesel energy content revealed that rapeseed and soybean-based biodiesels cannot be distinguished from one another, but that the animal-based biodiesels can be distinguished from the plant-based biodiesels (at a 99% confidence level).

In order to estimate fuel economy impacts using energy content, we needed a sampling of energy contents from conventional diesel fuel. Unfortunately, most fuel surveys available to us did not include measurements of energy content. We found only two sources from which to draw energy contents, shown in Table IV.A.2.a-2.

Table IV.A.2.a-2
Sources for diesel fuel energy content

	Average net Btu/gal
Database on which EPA Staff Discussion Document (July 2001) analyses were based	130,256
The Transportation Data Energy Book., Edition 17, 1997	128,700
Average	129,500

The first source provided 111 separate measurements of fuel energy content, but some of these fuels might not be the best representation of in-use diesel fuels because they were designed for use in engine test programs. The second source presented only the average for some unspecified sampling of conventional fuels. Without a straightforward means for evaluating the relative accuracy of the two estimates, we decided to use the average Btu/gal from the two sources.

Figure IV.A.2.a-1 shows how the distribution of volumetric energy contents differs for biodiesel and conventional diesel. The difference in energy content between pure biodiesel and conventional diesel fuel is summarized in Table IV.A.2.a-3.

Figure IV.A.2.a-1
Distribution of net energy content for biodiesel and conventional diesel

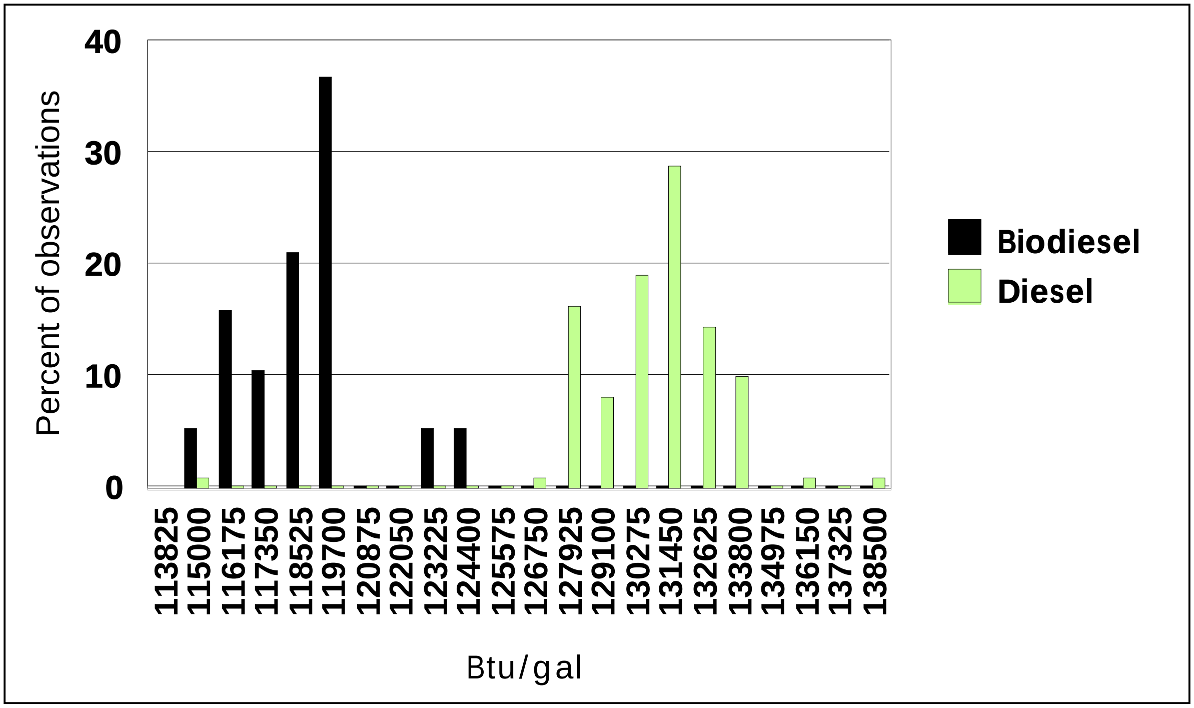


Table IV.A.2.a-3.
Difference in energy content between biodiesel and conventional diesel fuel

Conventional diesel	129,500 Btu/gal
Animal-based biodiesel	115,720 Btu/gal
Percent difference	-10.6 %

Conventional diesel	129,500 Btu/gal
Plant-based biodiesel	119,216 Btu/gal
Percent difference	-7.9 %

Since the energy content of two fuels is an excellent predictor of the relative fuel economy and volumetric energy content is expected to blend linearly, the percent differences in Table IV.A.2.a-3 can be applied directly to biodiesel blend fuel economy. Thus a 20 vol% biodiesel blend would be expected to exhibit a 2.1% (10.6% × 20%) reduction in fuel economy relative to the base fuel if the biodiesel in question was produced from animal fats, and a 1.6% reduction in fuel economy if the biodiesel was produced from soybeans or rapeseeds.

b. Via correlations with fuel consumption

The number of observations in our database for brake-specific fuel consumption (BSFC) was far less than that for the primary pollutants. For instance, the total number of NOx observations for heavy-duty highway engines that we used in our analysis was 626, while the number of available fuel consumption observations was only 217. As a result, we did not investigate all of the various adjustment terms for fuel consumption that we did for the pollutant correlations. Based on our work with the pollutant correlations, we concluded a priori that a squared biodiesel concentration term and adjustments for test cycle should not be included. Also, since adjustments for engine standards groups were largely unnecessary in the pollutant correlations (see Section IV.B.3), we did not investigate them for the fuel consumption correlation.

We compared the number of biodiesel observations for fuel consumption to our minimum data criteria to determine what other adjustment terms should be investigated. This analysis showed that the only adjustment term that could be investigated was rapeseed-based biodiesel source.

As for the individual pollutant correlations, we used proc_mix in SAS to determine how biodiesel concentration varied with the natural log of brake-specific fuel consumption (in units of lb/bhp-hr). This analysis indicated that the rapeseed adjustment term was not significant, so it was dropped and the correlation regenerated. The resulting correlation is shown below:

$$\text{BSFC, lb/hp-hr} = \exp[0.0008189 \times (\text{vol\% biodiesel}) - 0.855578]$$

This correlation indicates that fuel consumption increases as biodiesel is added to conventional diesel fuel. This result is consistent with the observation in Section IV.A.2.a above that fuel economy (proportional to the reciprocal of fuel consumption) decreases with increasing biodiesel concentration.

In order to make a direct comparison between the above correlation for mass-based fuel consumption and the volume-based fuel economy discussed in Section IV.A.2.a, we made several modifications to the BSFC correlation. These included taking the inverse of the BSFC correlation, adding a mass-to-volume units conversion using the specific gravities given in Table II.E.1-1, and then converting the equation into a percent change format. The result is shown below:

$$\begin{aligned}\text{\% change in fuel economy} = \Big\{ & \{\exp[-\ 0.0008189 \times (\text{vol\% biodiesel})] \\ & \times [0.88 \times (\text{vol\% biodiesel}/100) + 0.85 \times (1 - \text{vol\% biodiesel}/100)] \\ & \div 0.85\} - 1 \Big\} \times 100\%\end{aligned}$$

This version of the correlation between biodiesel concentration and percent change in fuel economy predicts that 100% biodiesel will result in a 4.6% reduction in fuel economy. This estimate is somewhat smaller than the values in Table IV.A.2.a-3. Since the 4.6% estimate is based on actual measurements of fuel consumption whereas the values in Table IV.A.2.a-3 are theoretical estimates based on energy content, we might have reason to place more confidence in the 4.6% estimate. However, the fact that it was necessary to make several modifications to the BSFC correlation introduces some additional uncertainty into the 4.6% estimate. As a result, we feel confident only in the range of possible impacts that biodiesel may have on fuel economy. These ranges are shown in Table IV.A.2.b-1.

Table IV.A.2.b-1
Fuel economy impacts of biodiesel use

	% reduction in miles/gallon
20% biodiesel	0.9 - 2.1
100% biodiesel	4.6 - 10.6

3. CO_2 impacts of biodiesel use

Since biodiesel is produced from plant oils or animal fats, it has been promoted as a means for reducing emissions of carbon dioxide that would otherwise be produced from the combustion of petroleum-based fuels. Carbon dioxide is considered by many to be an important component in global warming, though other pollutants can also play a role. The total impact that biodiesel could have on global warming would be a function not just of its combustion products, but also of the emissions associated with the full biodiesel production and consumption lifecycle. A study of the full lifecycle emission impacts of biodiesel is beyond the scope of this analysis, and we defer to previous and ongoing studies on this issue[6,7]. For this analysis, we have focused only on the calculation of CO_2 emissions as a function of biodiesel end use.

The carbon dioxide emissions impacts of biodiesel use was analyzed in much the same way that the regulated pollutants were analyzed. However, rather than first investigating each set of adjustment terms separately (base fuels, biodiesel source, and engine standards groups as discussed in Section IV.B below), we opted to only create a single correlation that examined them all at once. This approach is no different than that taken at the end of our assessment of regulated pollutants, described in Section IV.B.6 below. We also made the assumption, based on our analyses of regulated pollutants, that test cycle effects and engine standards group effects could be safely ignored during the analysis of carbon dioxide.

Before deciding which adjustment terms to include in our CO_2 correlation, we first examined the various categories of data and compared them to our minimum data criteria. As a result we determined that adjustment terms could be investigated for clean base fuels, animal-based biodiesel, and rapeseed-based biodiesel. In keeping with our approach to analyzing regulated pollutant effects, we used a backwards elimination process to settle on a set of statistically significant terms. The final correlation for CO_2 is shown below:

% change in CO_2 =
$\{\exp[\ + 0.0000177 \quad \times \text{(vol\% biodiesel)}$

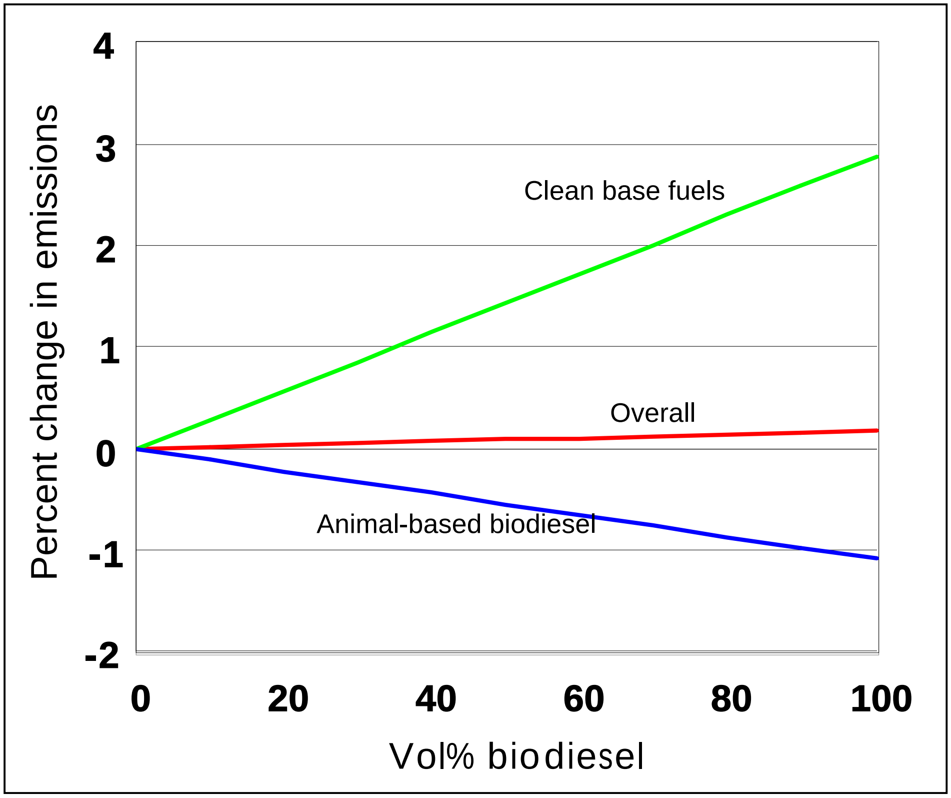
$$\begin{aligned} &+ 0.0002664 \times \mathrm{CLEAN} \times (\mathrm{vol\%\ biodiesel}) \\ &- 0.0001266 \times \mathrm{ANIMAL} \times (\mathrm{vol\%\ biodiesel}) \quad] - 1\} \times 100\% \end{aligned}$$

where

vol% biodiesel = Value from 0 to 100
CLEAN = 1 if the base fuel meets the conditions for "Clean" fuel given in Table III.C.2.e-1; otherwise, CLEAN = 0
ANIMAL = 1 if the biodiesel is produced from animal fat, tallow, or lard as described in Section III.C.2.c; otherwise, ANIMAL = 0

Note that the the overall vol% biodiesel term represents all biodiesel which is not animal-based and which has not been added to a 'clean' base fuel. Thus the overall vol% biodiesel term actually represents plant-based biodiesel added to an 'average' base fuel. The coefficient for this overall vol% biodiesel term was not statistically significant, but we retained it in the above equation to maintain hierarchy with the adjustment terms. Even so, the result is that this correlation predicts almost no effect of biodiesel on emissions of carbon dioxide for plant-based biodiesel added to an 'average' base fuel. The correlation does predict that the addition of biodiesel to a clean base fuel causes a moderate increase in carbon dioxide emissions, while the use of animal-based biodiesel with an average base fuel is predicted to produce a slight decrease. These results are shown graphically in Figure IV.A.3-1.

Figure IV.A.3-1
Biodiesel impacts on CO_2 emissions



The correlations suggest that biodiesel will produce higher CO_2 emissions if it is blended with a clean base fuel instead of an average base fuel. This result appears to be inconsistent with the correlations for HC and CO (see Figures IV.B.4-3 and IV.B.4-4), in that higher emissions of HC and CO should be associated with lower, not higher, emissions of CO_2. Similarly, the fact that animal-based biodiesel is predicted to produce less CO (see Figure IV.B.2-3) than plant-based biodiesel would suggest that animal-based biodiesel should produce higher, not lower, CO_2 emissions than plant-based biodiesel. However, the difference in CO_2 emissions impacts for animal and plant-based biodiesel might also be the result of different carbon contents for these two types of biodiesel. Therefore, we also examined the available data on carbon content of biodiesel versus that of conventional diesel to determine if it suggested differences in total carbon emitted.

Our database contained measures of H/C ratio and oxygen content for seventeen 100% biodiesel fuels. From this data we calculated the carbon content in terms of wt% carbon. The results are shown in Table IV.A.3-1, along with an estimate of the carbon content of conventional diesel fuel (0% biodiesel) drawn from the same database.

Table IV.A.3-1
Carbon content of biodiesel and conventional diesel (w/w %)

Biodiesel	
All biodiesel	77.3 %
Plant-based	77.8 %
Animal-based	76.1 %
Conventional diesel	86.7 %

The difference between plant and animal-based biodiesel carbon content is small (approximately 2%) but statistically significant at the $p = 0.1$ level. This difference is directionally consistent with that observed in Figure IV.A.3-1, where the "Overall" curve represents all plant-based biodiesel.

We can convert the wt% carbon values from Table IV.A.3-1 into pounds of carbon per gallon of fuel by multiplying them by the density (see Table II.E.1-1). Table IV.A.3-2 shows the results.

Table IV.A.3-2
Carbon content of biodiesel and conventional diesel (lb carbon / gal fuel)

Biodiesel	
All biodiesel	5.69
Plant-based	5.75
Animal-based	5.57
Conventional diesel	6.15

Although it would appear from the values in Table IV.A.3-2 that biodiesel would have a tendency to produce less CO_2 emissions than conventional diesel fuel, biodiesel also contains less energy per gallon. Thus more biodiesel must be consumed to propel a diesel vehicle a given distance than conventional diesel fuel. In order to take the difference in energy content into account, we calculated the carbon content per Btu by dividing the values in Table IV.A.3-2 by the volumetric energy contents given in Tables IV.A.2.a-1 and IV.A.2.a-2. The results are shown in Table IV.A.3-3.

Table IV.A.3-3
Carbon content of biodiesel and conventional diesel (lb carbon / million Btu)

Biodiesel	
All biodiesel	48.1
Plant-based	48.2
Animal-based	48.1
Conventional diesel	47.5

On an energy basis, the distinction between animal and plant-based biodiesel disappears. It also appears that biodiesel may actually increase emissions of CO_2 relative to conventional diesel fuel. However, this potential increase is small (~ 1%), and it is unlikely to be discernable in-use given the variability in each of the components (density, H/C ratio, and energy content). These results suggest that there would likely be no measurable difference between biodiesel and conventional diesel in terms of exhaust CO_2 emissions. The CO_2 correlations shown in Figure IV.A.3-1 suggest that a difference between biodiesel and conventional diesel may exist, but those correlations were also somewhat inconsistent with the correlations of HC and CO as described above. Therefore, we cannot confidently conclude that biodiesel increases or decreases CO_2 emissions based on these analyses. This issue warrants further investigation, as the impacts of biodiesel on CO_2 inventories is one reason for the heightened interest in biodiesel in recent years.

B. Investigation of adjustment terms for regulated pollutants

Because there were a number of different categories of potential adjustment terms, we decided first to investigate each separately. Once we determined which adjustment terms were important, we then developed a strategy for incorporating them all into a single correlation for each pollutant. Each of the subsections below describes our results for the separate adjustment term investigations, and Section IV.B.6 describes our composite correlations.

1. Test cycle effects

For our initial analysis of heavy-duty highway engines, we set aside all data collected on generic transient or steady-state cycles. This left only composite FTP, hot-start FTP, and R49 data. We introduced terms into the regression analysis representing the interaction of these test cycle categorical variables with percent biodiesel to determine which ones might be statistically significant. As described in Section III.C.2.b, coefficients could be estimated for only two of the three potential interactive terms because an overall % biodiesel term was also included. The p-values for the % biodiesel terms are listed in Table IV.B.1-1, while the estimated coefficients are given in Table IV.B.1-2.

Table IV.B.1-1
P-values for test cycle effects

	NOx	PM	HC	CO
% biodiesel	**0.0002**	**0.0001**	**0.0001**	**0.0001**
R49 × % biodiesel	0.4144	**0.0256**	0.3733	**0.0016**
UDDS × % biodiesel	0.1782	0.7031	0.5709	0.2358
UDDSH × % biodiesel	NA	NA	NA	NA

Table IV.B.1-2
Estimated coefficients for test cycle effects

	NOx	PM	HC	CO
% biodiesel	0.000855	-0.006292	-0.011469	-0.006863
R49 × % biodiesel	0.000600	0.005147	0.006599	0.005646
UDDS × % biodiesel	0.000131	-0.000151	-0.000654	0.000335
UDDSH × % biodiesel	0.000000	0.000000	0.000000	0.000000

For the test cycle adjustment terms, p-values greater than 0.05 suggest that the effect in question cannot be confidently distinguished from zero. In these cases the statistically significant overall '% biodiesel' term would apply to the individual test cycle term as well, indicating that there is no variation in the biodiesel effect on emissions as a function of test cycle. The two exceptions to this conclusion are for PM and CO where R49 data appears to exhibit different biodiesel effects than data collected on the FTP. This result is consistent with previous experience: steady-state test cycles are generally not accurate predictors of the PM and CO emissions that would be generated under transient conditions. Based on this result, we decided to exclude R49 data from our final PM and CO correlations.

2. Biodiesel source effects

All biodiesel blends in our database were placed into one of three biodiesel source categories: soybean, rapeseed/canola, and animal. We introduced terms into the regression analysis which represented the interaction of these biodiesel source categorical variables with percent biodiesel to determine which ones might be statistically significant. As described in Section III.C.2.c, coefficients could be estimated for only two of the three potential interactive terms because an overall % biodiesel term was also included. The p-values for the % biodiesel terms are listed in Table IV.B.2-1, and the estimated coefficients are given in Table IV.B.2-2. P-values lower than 0.05 were considered significant.

Table IV.B.2-1
P-values for biodiesel source effects

	NOx	PM	HC	CO
% biodiesel	**0.0001**	**0.0001**	**0.0001**	**0.0003**
Animal × % biodiesel	**0.0001**	**0.0001**	0.5525	**0.0001**
Rape × % biodiesel	**0.0311**	0.6316	0.9162	**0.0164**
Soy × % biodiesel	NA	NA	NA	NA

Table IV.B.2-2
Estimated coefficients for biodiesel source effects

	NOx	PM	HC	CO
% biodiesel	0.001553	-0.000908	-0.001031	-0.000603
Animal × % biodiesel	-0.001216	0.024410	0.000022	-0.000838
Rape × % biodiesel	-0.000331	0.007517	0.000038	-0.000316
Soy × % biodiesel	0.000000	0.000000	0.000000	0.000000

For NOx and CO, the three biodiesel source categories appear to produce three different correlations between biodiesel concentration and emissions. For PM, animal-based biodiesel appears to differ from plant-based biodiesel. For HC, there is no discernable difference between the three source categories. The biodiesel source effects for NOx, PM, and CO are shown in Figures IV.B.2-1, IV.B.2-2, and IV.B.2-3, respectively.

Figure IV.B.2-1
Biodiesel source effects for NOx

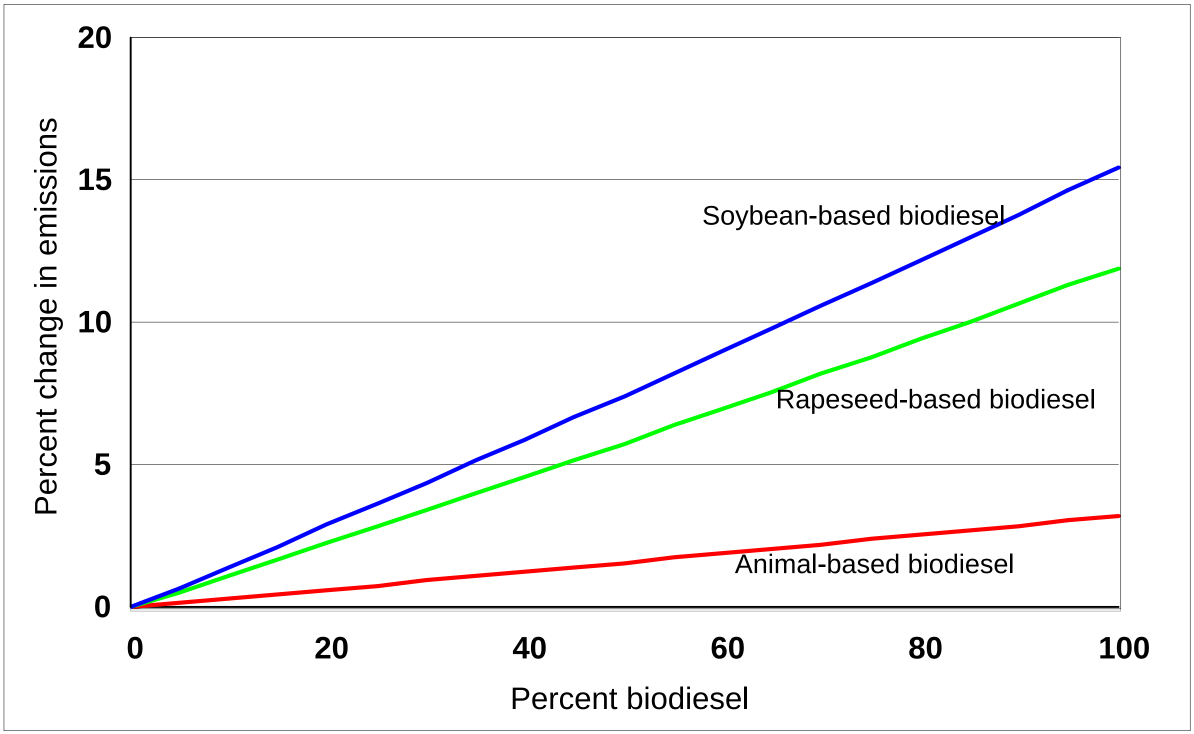


Figure IV.B.2-2
Biodiesel source effects for PM

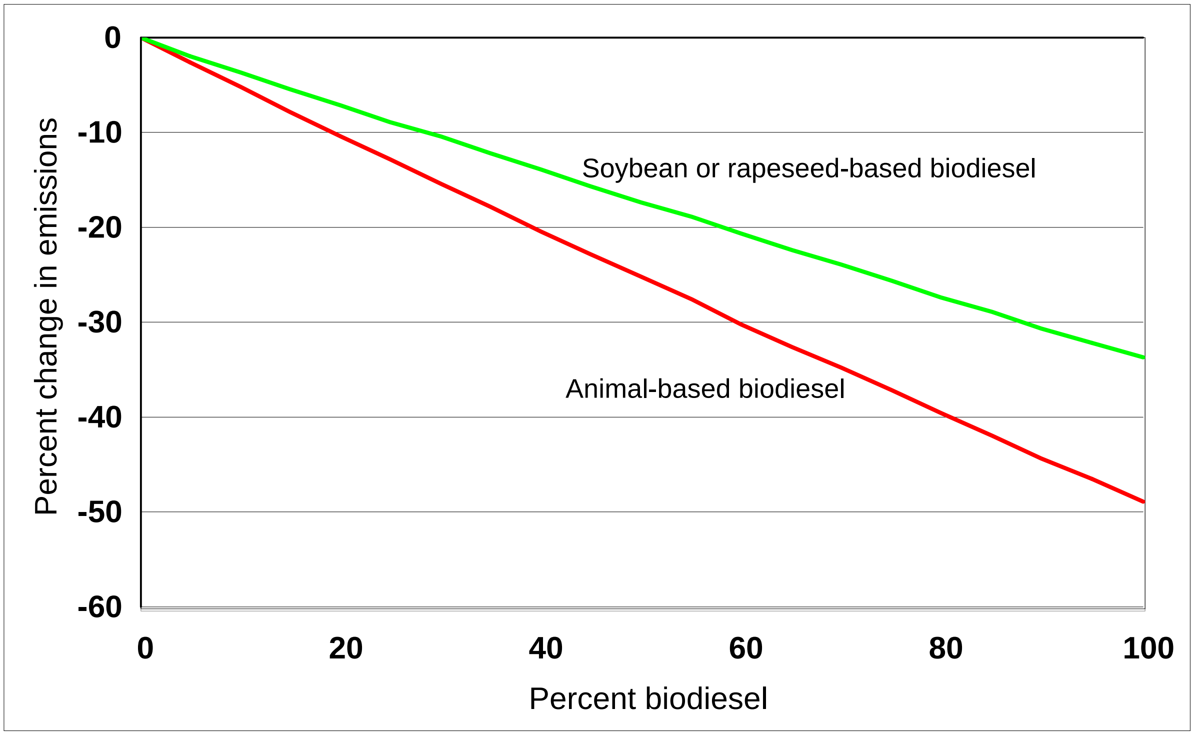
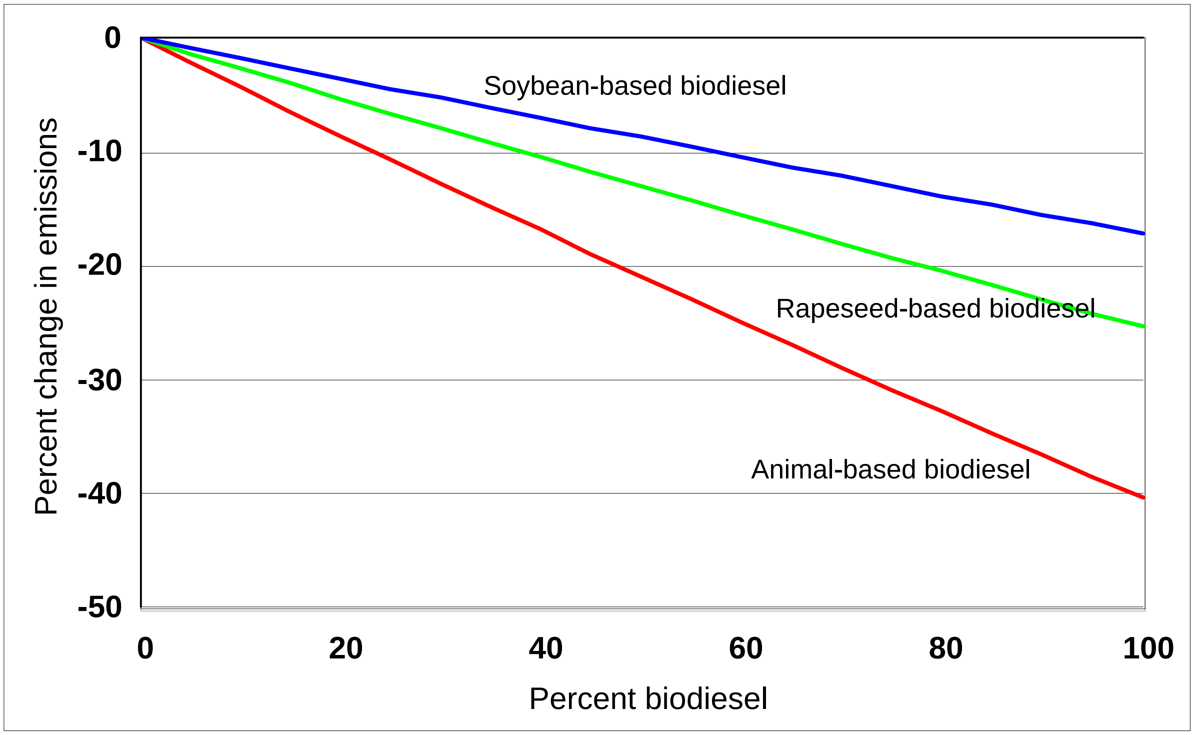


Figure IV.B.2-3
Biodiesel source effects for CO



3. Engine standards groups

All biodiesel blends in our database were placed into one of several engine standards groups. Of these, only three groups met our minimum data criteria: D, E, and G. We therefore introduced terms into the regression analysis which represented the interaction of these three engine standards groups (as categorical variables) with percent biodiesel to determine which ones might be statistically significant. The p-values for the % biodiesel terms are listed in Table IV.B.3-1, and the estimated coefficients are given in Table IV.B.3-2. P-values lower than 0.05 were considered significant.

Table IV.B.3-1
P-values for engine standards groups

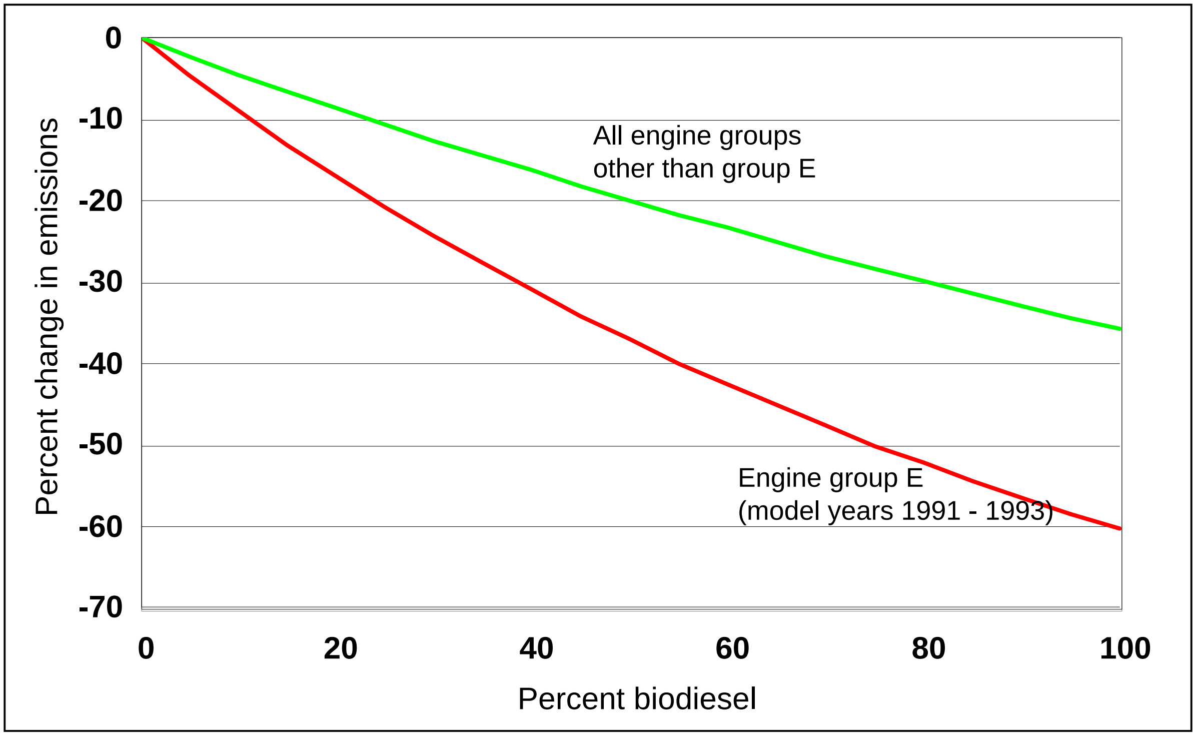
	NOx	PM	HC	CO
% biodiesel	**0.0074**	0.4005	0.2939	**0.0066**
Group D × % biodiesel	0.2598	0.3008	0.3383	0.5056
Group E × % biodiesel	0.1678	**0.0147**	0.3407	0.0676
Group G × % biodiesel	0.4056	0.6825	0.5549	0.2035

Table IV.B.3-2
Estimated coefficients for engine standards groups

	NOx	PM	HC	CO
% biodiesel	0.001382	-0.002260	-0.006038	-0.004081
Group D × % biodiesel	-0.000720	-0.002979	-0.006813	-0.001235
Group E × % biodiesel	-0.000844	-0.006973	-0.006515	-0.003257
Group G × % biodiesel	0.000575	-0.001240	-0.004582	-0.002567

There is only one case in which an engine standards group interactive term is statistically significant: Standards Group E for PM, representing model years 1991 through 1993. In this case, the adjustment terms for Groups D and G were dropped and the correlation regenerated, after which both the overall '% biodiesel' term and the Group E adjustment terms were statistically significant. Figure IV.B.3-1 shows the effect of the Group E adjustment term on PM.

Figure IV.B.3-1
Engine standard group effects for PM



4. Base fuel effects

All biodiesel blends in our database were placed into one of two base fuel groups: clean or average. We therefore introduced terms into the regression analysis representing the interaction of these two categorical base fuel groups with percent biodiesel to determine which ones might be statistically significant. As described in Section III.C.2.e, coefficients could be estimated for only one of the two potential interactive terms because an overall % biodiesel term was also included. The p-values for the % biodiesel terms are listed in Table IV.B.4-1, and the estimated coefficients are given in Table IV.B.4-2. P-values lower than 0.05 were considered significant.

Table IV.B.4-1
P-values for base fuel groups

	NOx	PM	HC	CO
% biodiesel	**0.0001**	**0.0001**	**0.0048**	**0.0001**
Average × % biodiesel	**0.0001**	**0.0004**	**0.0015**	**0.0001**
Clean × % biodiesel	NA	NA	NA	NA

Table IV.B.4-2
Estimated coefficients for base fuel groups

	NOx	PM	HC	CO
% biodiesel	0.002467	-0.004355	-0.007087	-0.004544
Average × % biodiesel	-0.001706	-0.001753	-0.004757	-0.001820
Clean × % biodiesel	0.000000	0.000000	0.000000	0.000000

For all four pollutants, the base fuel appears to have a significant impact on the correlation between biodiesel concentration and emissions. This result suggests that the base fuel effects on the emission impacts of biodiesel blending should be investigated further. The base fuel effects for NOx, PM, HC, and CO are shown in Figures IV.B.4-1, IV.B.4-2, IV.B.4-3, and IV.B.4-4 respectively.

Figure IV.B.4-1
Base fuel effects for NOx

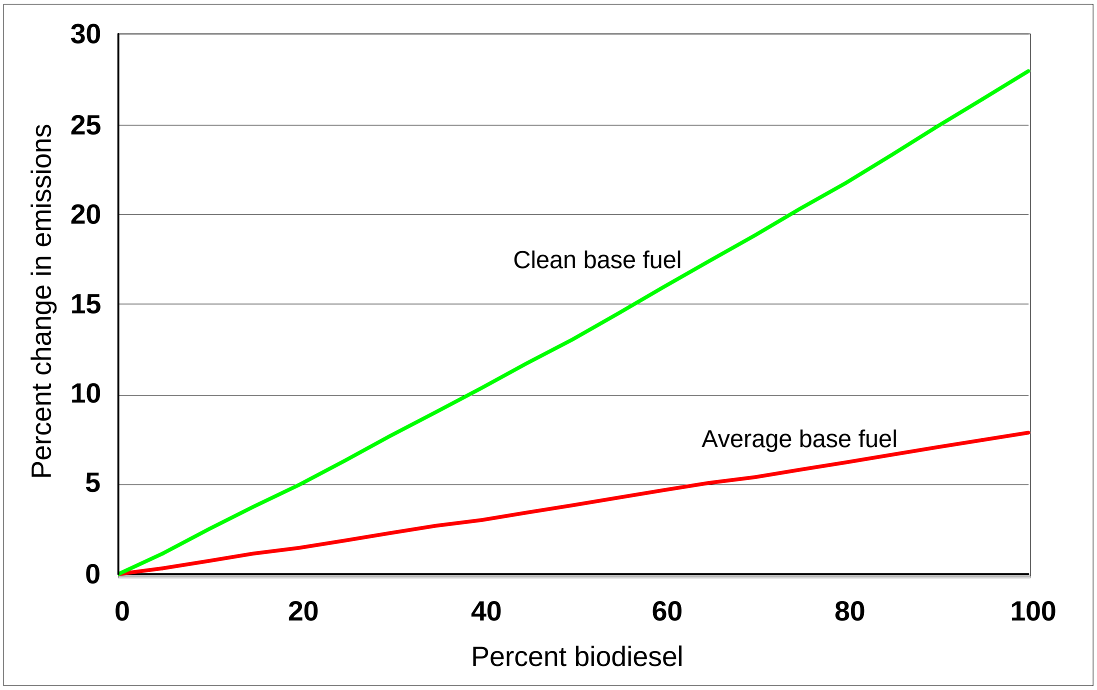


Figure IV.B.4-2
Base fuel effects for PM

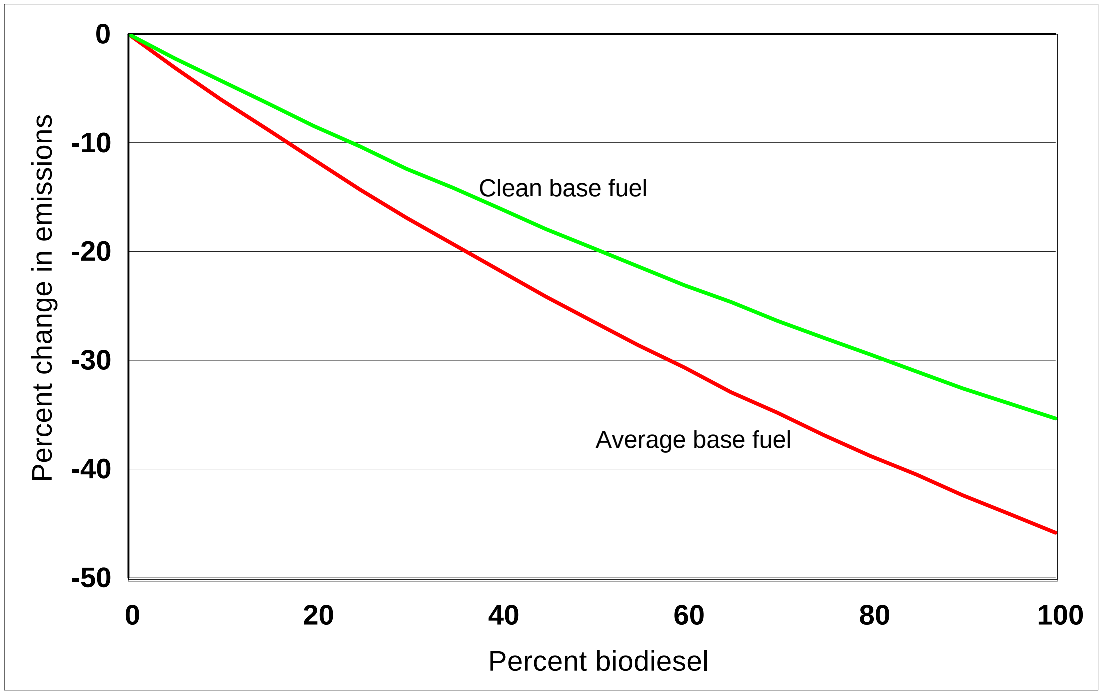


Figure IV.B.4-3
Base fuel effects for HC

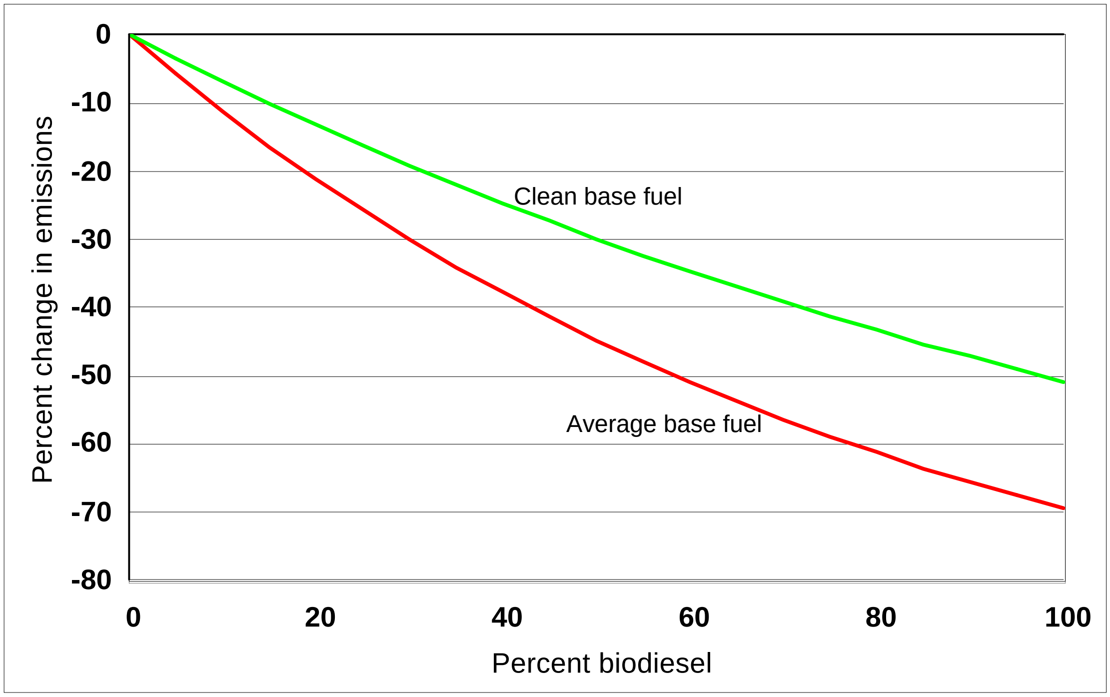
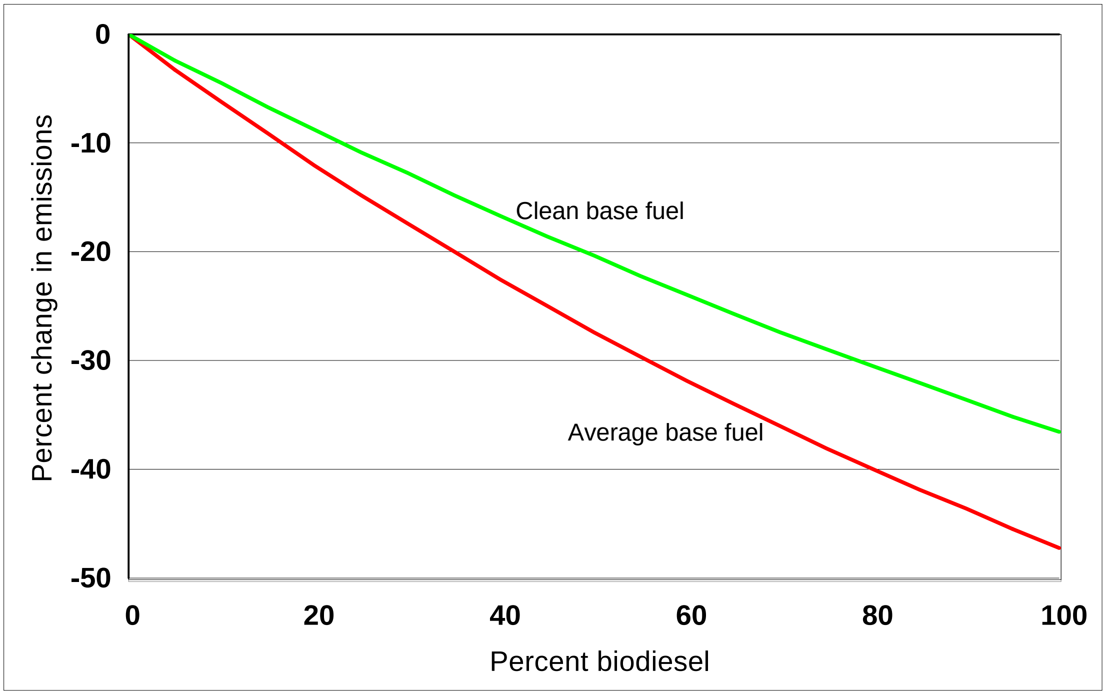


Figure IV.B.4-4
Base fuel effects for CO



5. Cetane effects

We conducted an analysis to determine the importance of cetane in the correlation of biodiesel concentration with NOx emissions. This analysis was not done in order to add a cetane term to the basic correlations in Section IV.A, but rather to determine if it was legitimate to extrapolate the effects of the basic correlations to EGR-equipped engines as described in Section III.C.2.f.

In order to determine the degree to which cetane plays a role in the biodiesel/NOx relationship, we separated cetane effects in the biodiesel database from the measured NOx emissions. To do this, we used the following equation to predict the NOx emissions for every fuel in our biodiesel database,

$$\mathrm{NOx} = \exp(-0.004512 \times (\mathrm{CETNUM} - 45) + 0.0001458 \times (\mathrm{CETNUM} - 45)^2 + 1.5497034)$$

and then calculated the difference between the predicted NOx and measured NOx for every observation. These "pseudo-residuals" were therefore independent of any cetane effects, and could be used as dependent variables in a mixed model analysis of biodiesel concentration. Depending on the magnitude of the biodiesel effect for this correlation and its associated statistical significance, we could determine the degree to which cetane plays a role in the relationship between biodiesel concentration and NOx emissions.

Table IV.B.5-1 gives the results of this analysis. Since the correlation made use of only those observations for which the cetane number was available, the results are not directly comparable to the basic NOx correlation presented in Section IV.A. Therefore, for comparison, we reconstructed the basic NOx correlation using only observations that included a cetane number measurement. This modified basic NOx correlation is also shown in Table IV.B.5-1.

Table IV.B.5-1
Mixed model results for cetane analysis

Dependent variable	Coefficient for % biodiesel	Intercept	P-value for % biodiesel
ln(pseudo-residuals + 1.1[α])	0.001027	0.474478	0.5900
ln(NOx)	0.000694	1.734796	0.0008

[α] The factor 1.1 was added to ensure that the natural log was always defined

The pseudo-residuals do not appear to be correlated with biodiesel concentration, given the high P-value. This result suggests that cetane number is highly correlated with biodiesel concentration, so that the removal of cetane effects on NOx from the NOx measurements leaves little left to be explained by biodiesel concentration.

This result does not imply that cetane number is a better predictor of biodiesel effects on NOx than biodiesel concentration. Rather, it highlights a collinearity between the two, a collinearity that could be inferred from the different average cetane numbers between biodiesel and conventional diesel as shown in Table II.E.1-1. As a result of this collinearity, a wide variety of correlations having equivalent explanatory power could be developed that have either cetane number, biodiesel concentration, or some combination of the two. For our purposes, correlations based strictly on biodiesel concentration are more practical because those who use biodiesel don't always have a measurement of its cetane number. Note that the cetane number of biodiesel does appear to differ between plant and animal-based biodiesel, as described in Section II.E.1. Thus the cetane number of biodiesel is being taken into account in an indirect fashion through the investigation of biodiesel source effects and base fuel effects, as described in Sections IV.B.2 and IV.B.4, respectively.

The analysis summarized in Table IV.B.5-1 indicates that cetane number is an important component of the impact that biodiesel has on NOx emissions, and we know from previous analyses that EGR-equipped engines do not respond in any significant way to changes in cetane. Therefore, it might seem reasonable not to apply the fleet-average biodiesel effects on NOx to EGR-equipped engines. In practical terms, this would mean that our NOx correlations would not apply to model year 2002 and later engines (engine standards group B, per Table II.C-1). However, this approach would actually mean that the NOx detriment associated with biodiesel would diminish as the fleet turns over. Since we do not have any actual data showing biodiesel effects on EGR-equipped engines, we cannot verify that this result is accurate, and thus there is a possibility that this approach would underestimate the true NOx impact of biodiesel for future fleets. Therefore, we have determined that the more environmentally conservative approach is to continue to allow our NOx correlations to apply to the entire fleet, include EGR-equipped engines. We welcome comment on this approach and any alternative approaches for dealing with the lack of biodiesel data on EGR-equipped engines.

6. Composite correlations

The analyses described above suggest that the correlation between biodiesel concentration and emissions should include adjustments for biodiesel source, engine standards groups, and base fuel. We therefore developed a single correlation for each pollutant that included any adjustment terms that had been deemed important in the previous analyses.

We first noted that some potential adjustment terms were redundant, given that the correlations would all include an overall '% biodiesel' term. This fact is reflected in the 'NA' results in Tables IV.B.2-1 and IV.B.4-1. As a result we assigned a hierarchy to the adjustment terms within each category to help determine which adjustment terms should be investigated in the composite correlation. For base fuel effects, "Clean" was chosen as the adjustment term, so that the unadjusted '% biodiesel' term would necessarily reflect "Average" base fuels. For biodiesel source effects, "Animal" and "Rape" were chosen as the adjustment terms, so that the unadjusted '% biodiesel' term would necessarily reflect "Soy". Both of these choices reflect the most common circumstances under which biodiesel is currently used.

We then made a list of all adjustment terms that we believed should be introduced into each pollutant correlation, based on the previous analyses[e]. This list included not only the first-order interactions between % biodiesel and the individual categorical variables, but also second-order interactions between one categorical variable and another, in addition to their combined interaction with % biodiesel. These second-order interactions needed to be considered because the previous analyses provided no information on whether one type of adjustment is affected by the presence of another type of adjustment. Table IV.B.6-1 lists the terms that we intended to investigate in our composite correlation, based on previous analyses.

Table IV.B.6-1
Adjustment terms to consider in composite correlations

	NOx	PM	HC	CO
First-order	Clean Animal Rape	Clean Animal Group E	Clean	Clean Animal Rape
Second-order	Clean × Animal Clean × Rape	Clean × Animal Clean × Group E Animal × Group E		Clean × Animal Clean × Rape

The next step was to determine if there was sufficient data in the database to investigate all of the adjustment terms listed in Table IV.B.6-1. We therefore counted all heavy-duty highway engine biodiesel observations for every pollutant and compared them to our minimum data criteria. We determined that there was insufficient data to investigate any of the second-order terms involving biodiesel source (i.e. Animal or Rape) and base fuel (i.e. Clean), but that all of the first-order terms had sufficient data to be investigated.

We also discovered that there were several cases in which the same data could be used to produce either a first order adjustment term or a second-order adjustment term involving engine standards group E (model years 1991 to 1993). In these cases, all the relevant data was collected on engines falling into engine standards group E. These redundant terms are listed in Table IV.B.6-2.

---

[e] We also conducted an all-possible-terms analysis using stepwise backwards elimination to verify that the final selection of correlation terms was appropriate. This additional analysis is described at the end of this Section.

Table IV.B.6-2
Redundant adjustment terms in composite correlation

NOx	(Animal) and (Animal × Group E)
PM	(Clean) and (Clean × Group E) (Animal) and (Animal × Group E)
HC	n/a
CO	(Clean) and (Clean × Group E) (Animal) and (Animal × Group E)

Since group E contains the most data of all engine standards groups, we could, for example, legitimately apply the Animal adjustment term in the correlation for NOx to the entire fleet or only to engine standards group E. We chose to use an environmentally conservative approach in which we made this decision based on whether the adjustment term in question would increase or decrease the emission effects of biodiesel. If one of the adjustment terms listed in Table IV.B.6-2 would result in increased emissions, we chose to apply it to all engine groups. If instead it would result in decreased emissions, we chose to apply it only to engine standards group E.

Given that our previous analyses showed that all the first-order adjustment terms listed in Table IV.B.6-1 were important, we decided to use an approach to developing a composite correlation that retained as many of these adjustment terms as possible. Thus we used a backwards elimination approach that began with a correlation having all the first-order adjustment terms listed in Table IV.B.6-1 and then eliminated statistically insignificant terms one by one until all remaining terms were significant. The result was that few of the terms were dropped. The resulting "composite" correlations are shown below:

% change in NOx =

$$\left\{\exp\left[\begin{aligned}&+0.0010375 \times (\text{vol\% biodiesel})\\&+0.0012289 \times \text{CLEAN} \times (\text{vol\% biodiesel})\\&-0.0002732 \times \text{RAPE} \times (\text{vol\% biodiesel})\\&-0.0009795 \times \text{ANIMAL} \times \text{GROUP E} \times (\text{vol\% biodiesel})\end{aligned}\right]-1\right\} \times 100\%$$

% change in PM =

$$\left\{\exp\left[\begin{aligned}&-0.0047395 \times (\text{vol\% biodiesel})\\&+0.0010742 \times \text{CLEAN} \times (\text{vol\% biodiesel})\\&-0.0045908 \times \text{GROUP E} \times (\text{vol\% biodiesel})\\&-0.0019343 \times \text{ANIMAL} \times \text{GROUP E} \times (\text{vol\% biodiesel})\end{aligned}\right]-1\right\} \times 100\%$$

% change in HC =

$$\left\{\exp\left[\begin{aligned}&-0.0118443 \times (\text{vol\% biodiesel})\\&+0.0047569 \times \text{CLEAN} \times (\text{vol\% biodiesel})\end{aligned}\right]-1\right\} \times 100\%$$

% change in CO =

$$\begin{aligned}\{\exp[\; & -0.0058238 \times (\text{vol\% biodiesel}) \\ & +0.0010853 \times \text{CLEAN} \times (\text{vol\% biodiesel}) \\ & +0.0017335 \times \text{RAPE} \times (\text{vol\% biodiesel}) \\ & -0.0017116 \times \text{ANIMAL} \times \text{GROUP E} \times (\text{vol\% biodiesel}) \;] - 1\} \times 100\%\end{aligned}$$

where

vol% biodiesel = Value from 0 to 100
CLEAN = 1 if the base fuel meets the conditions for "Clean" fuel given in Table III.C.2.e-1; otherwise, CLEAN = 0
ANIMAL = 1 if the biodiesel is produced from animal fat, tallow, or lard as described in Section III.C.2.c; otherwise, ANIMAL = 0
RAPE = 1 if the biodiesel is produced from rapeseed oil or canola oil, as described in Section III.C.2.c; otherwise, RAPE = 0
GROUP E = 1 if the highway engines being evaluated are model years 1991 to 1993; otherwise, GROUP E = 0

We also used an alternative approach to developing correlations between biodiesel concentration and emission impacts. In this alternative approach, we did not limit the candidate adjustment terms for each pollutant to those we identified as important in Sections IV.B.2, IV.B.3, and IV.B.4. Instead, we permitted any term, including second-order terms, to be included in the regression analysis if the available data for that term met our minimum data criteria. Table IV.B.6-3 lists all the terms that met our minimum data criteria and thus were included in this alternative analysis.

Table IV.B.6-3
All possible terms for composite correlations

NOx	PM	HC	CO
Clean Animal$^{\alpha}$ Rape Group D Group E Group G Rape × Group D Clean × Group E	Clean$^{\alpha}$ Animal$^{\alpha}$ Rape Group D Group E Group G Rape × Group D	Clean Animal$^{\alpha}$ Rape Group D Group E Group G Rape × Group D Clean × Group E	Clean$^{\alpha}$ Animal$^{\alpha}$ Rape Group D Group E Group G Rape × Group D

$^{\alpha}$ These terms can also be represented as an interaction with engine standards group E, since all the data in question was collected on group E engines

As before, we used a (manual) stepwise backwards elimination process in SAS procedure proc_mix to drop terms that were not significant. When all the remaining terms were significant, we stopped. As a result of this "all possible terms" analysis, the PM, HC, and CO correlations were identical to the final composite correlations given above. The "all possible terms" NOx correlation, however, included an additional adjustment term for Group G. Further review of the "all possible terms" correlation for NOx indicated that the inclusion of this Group G term lowered fleet-average estimated NOx emissions by an average of 0.2 percent. Thus not only are the estimated effects of the "all possible terms" NOx correlation very similar to those from our original analysis, but the original analysis predicts slightly larger increases in NOx due to the use of biodiesel. We have chosen to use the original final composite correlation for NOx presented above, since the slightly higher NOx emissions provide a more environmentally conservative estimate of biodiesel impacts.

C. Comparison of vehicle data to engine data

Our database contained heavy-duty highway data collected on both engines and vehicles. The amount of engine data far surpassed the amount of vehicle data, as shown in Table III.B-1. Given that engines and vehicles have the potential for exhibiting different emission effects, we opted to base our primary biodiesel correlations on engine data only. However, we wanted to compare the vehicle data to predictions made by our engine-based correlations to determine if the impact of biodiesel on emissions can be considered to be the same for vehicles and engines.

To make this comparison, we first converted all vehicle emissions data from g/mile to % change in emissions. This process involved identifying the base fuel to which biodiesel had been added for each vehicle/test cycle combination, and averaging any repeat measurements made on this base fuel. The % change values for each biodiesel test were then calculated with respect to the averaged emissions from the repeat base fuel measurements.

We used our composite correlations to estimate the % change in emissions of each pollutant for the specific biodiesel concentration associated with each vehicle test. We then compared the predicted and observed % change emission values in several different ways to determine if our engine-based correlations could be said to represent vehicle effects of biodiesel. These comparisons included graphical comparisons, paired t-tests of predicted versus observed % change values, and comparisons of residuals for the engine and vehicle data. Details of how these comparisons were done are summarized in Section V.

Figures IV.C-1 through IV.C-4 compare the predicted and observed values for each of the four pollutants.

Figure IV.C-1
Predicted versus observed NOx for vehicles

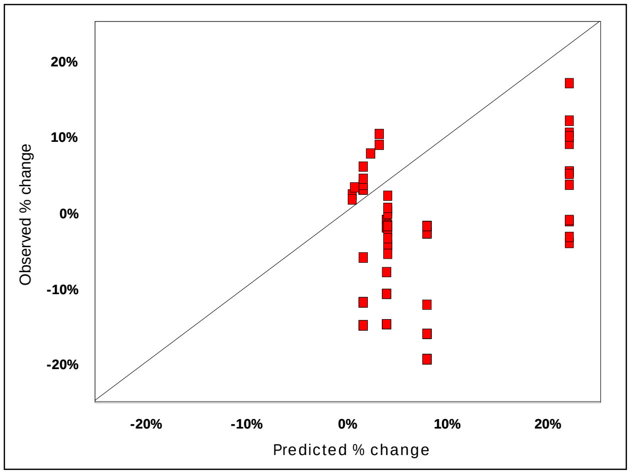


Figure IV.C-2
Predicted versus observed PM for vehicles

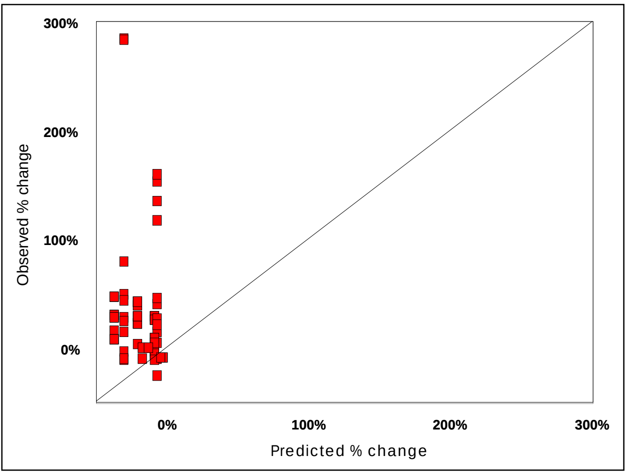


Figure IV.C-3
Predicted versus observed CO for vehicles

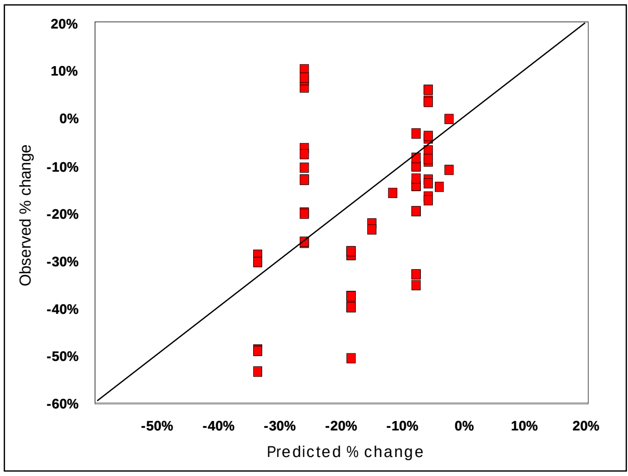
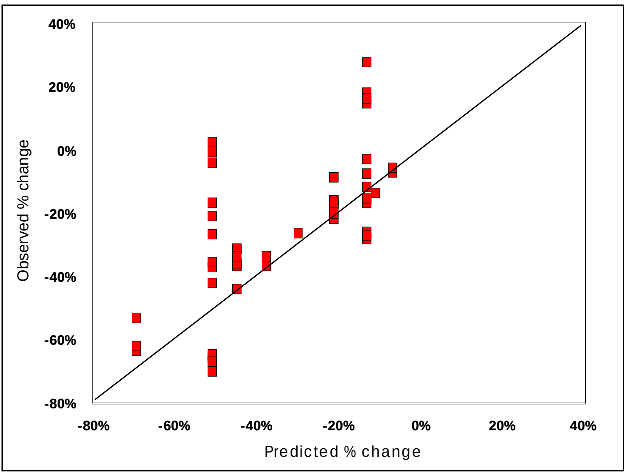


Figure IV.C-4
Predicted versus observed HC for vehicles



These figures suggest that the emission effects measured for heavy-duty vehicles exhibit a bias in comparison to the estimates predicted using our composite correlations. A t-test of predicted versus observed values confirms that the bias is statistically significant for all four pollutants. In addition, this bias is not consistent. For PM and HC, the vehicle data appears to produce emission benefits that are smaller than those predicted by the composite correlations. For NOx the vehicle data appears on average to produce emission reductions whereas the composite correlations predict emission increases. For CO, the vehicle data appears to produce larger emission benefits than the composite correlation predictions. Based on this comparison, we do not believe that the vehicle data can be used to represent the emission effects of biodiesel on heavy-duty diesel engines.

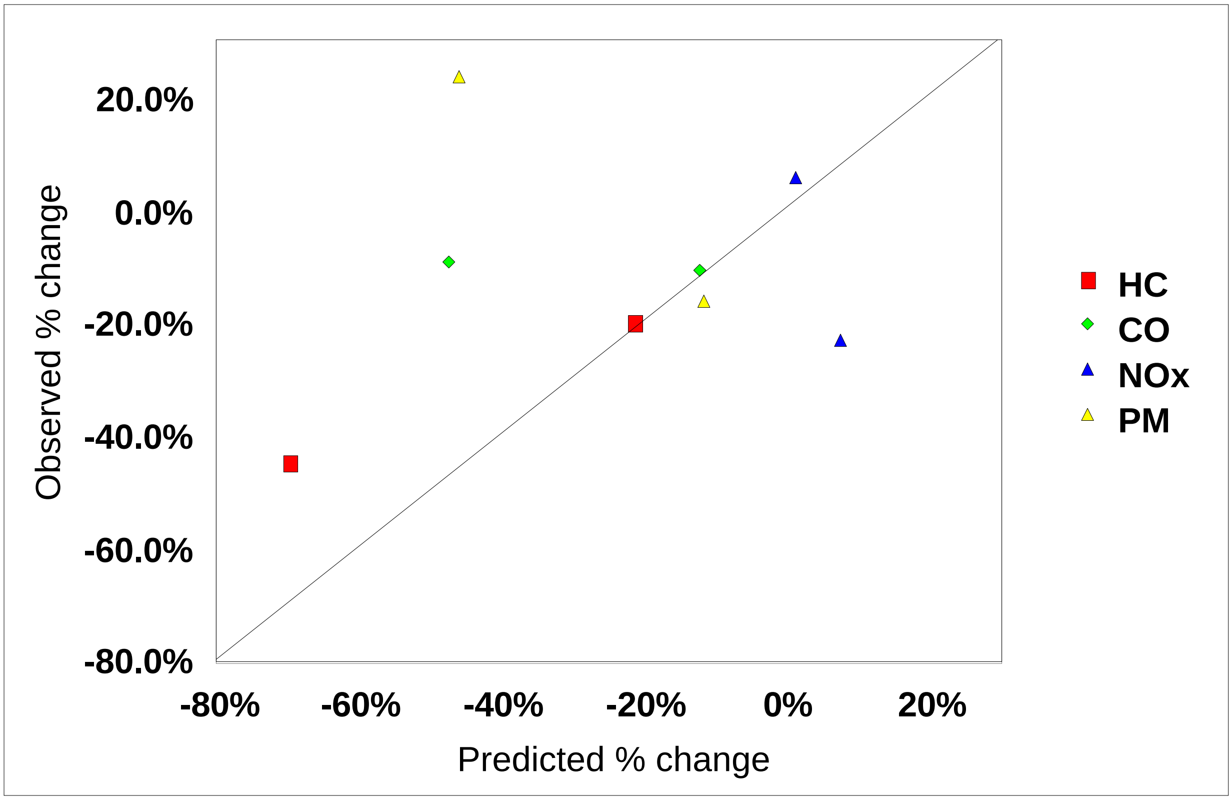
D. Use of virgin oils as biodiesel

In our analysis we excluded all biodiesels which were virgin oils rather than their transesterified counterparts. From the literature it appears that biodiesel research has focused primarily on esters due to their superior properties; virgin oils tend to have significantly lower cetane and higher viscosity than ester biodiesels and conventional diesel fuel. On this basis we would not expect virgin oils to play a significant role in biodiesel use in the future.

However, it seemed prudent to investigate whether the effects exhibited by our composite correlations would reflect the data on virgin oils. Unfortunately, our database contains only two unique measurements of virgin oil biodiesel collected on standard test cycles. We would not expect conclusions drawn on the basis of these two emission measurements to be determinate, but we offer the comparison of correlation predictions to virgin oil data nonetheless. Because the properties of virgin oils are so different from those of esters, we would not expect virgin oil biodiesel to produce the same emission effects as ester-based biodiesel. Unless a comparison of predicted and observed values provided strong evidence to the contrary, we would assume that our ester-based composite correlations would not be applicable to virgin oil biodiesels.

A graphical comparison does not suggest a 1:1 alignment of predicted and observed values (see Figure IV.D-1). As a result, we do not believe that our composite correlations should be applied to virgin oil biodiesels.

Figure IV.D-1
Comparison of virgin oils to ester-based biodiesel



E. Comparisons to other emission correlations

We believe that the analyses presented in this report represent the most comprehensive analyses done to date on the effects that biodiesel has on emissions of regulated pollutants. However, since we used all available and relevant data, essentially none is left with which to validate the correlation's predictive capabilities. We therefore compared our correlation's predictions to those from previous analyses conducted by other analysts.

We are aware of three different previous analyses that attempted to compile data from multiple test programs and draw conclusions regarding the impact that varying biodiesel concentrations have on emissions of regulated pollutants:

1. Howell, S., "Emissions Summary for Biodiesel," MARC-IV Consulting, Inc., memorandum to Sam McCahon, December 8, 2000
2. Lindhjem et al, "Impact of Biodiesel Fuels on Air Quality," ENVIRON International Corporation Task 1 Report, June 13, 2000
3. Sheehan et al, "Life Cycle Inventory of Biodiesel and Petroleum Diesel for Use in an Urban Bus," National Renewable Energy Laboratory Final Report, May 1998

These researchers made no attempt to account for such factors as test cycle, base fuel properties, biodiesel source effects, and engine technology. Therefore, a direct comparison of the correlations developed by the above researchers to our composite correlations may not be appropriate. Instead, we compared these other correlations to our basic correlations described in Section IV.A. Figures IV.E-1, IV.E-2, IV.E-3, and IV.E-4 provide this comparison. Note that the Lindhjem report includes average biodiesel impacts at only 20 vol% and 100 vol%. The associated emission impacts were linearly interpolated for the purposes of graphical presentation.

Figure IV.E-1
Comparison of basic EPA correlation to alternative correlations for NOx

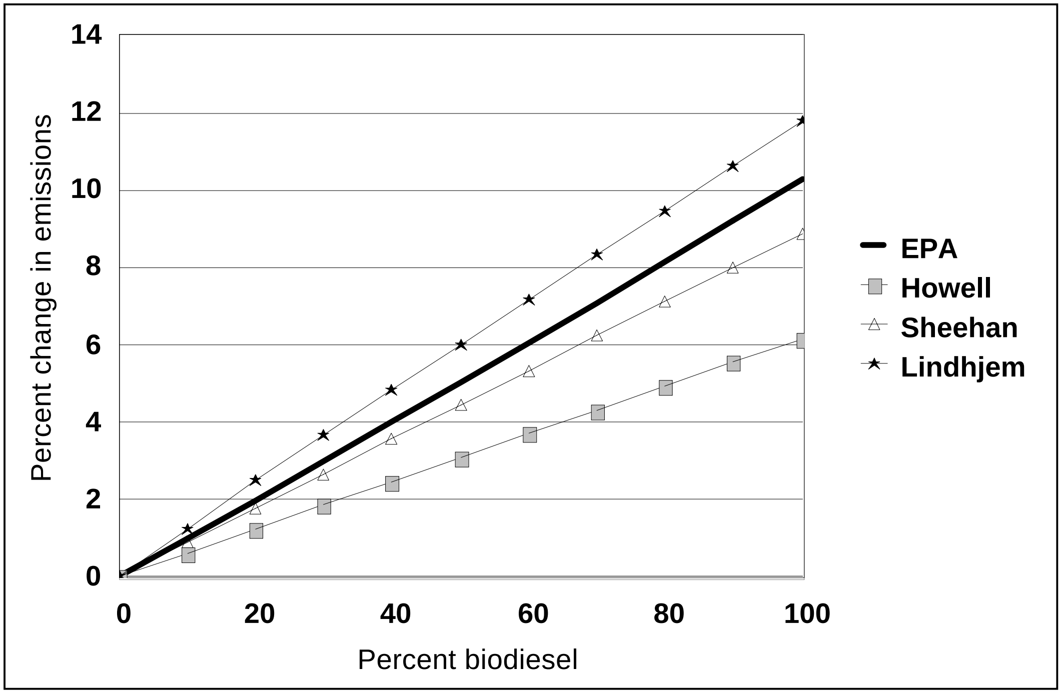


Figure IV.E-2
Comparison of basic EPA correlation to alternative correlations for PM

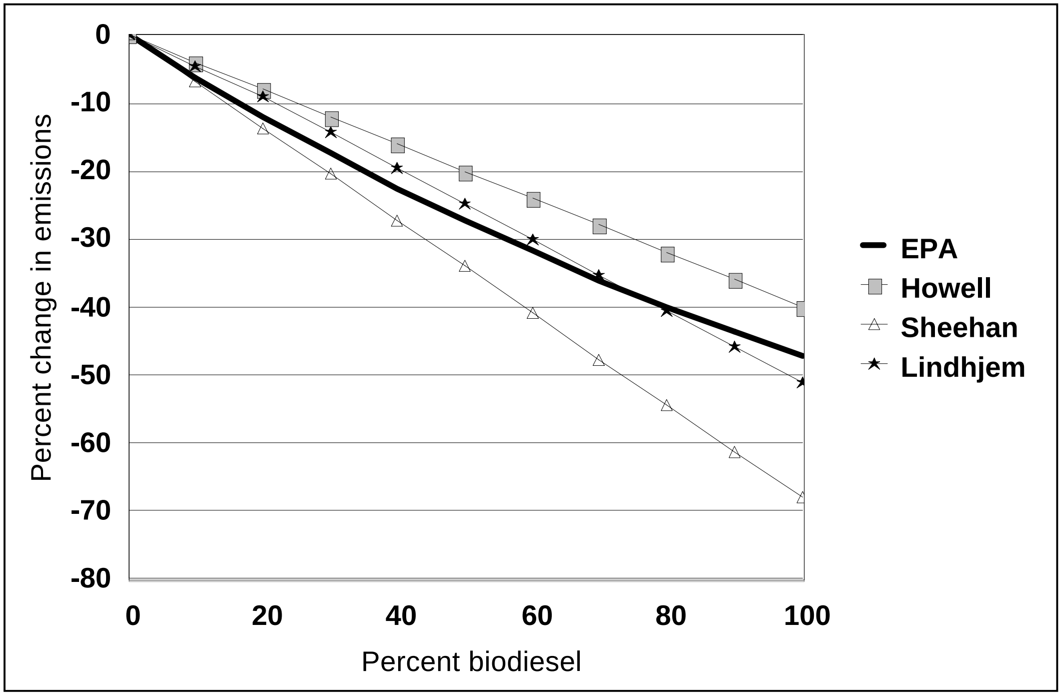


Figure IV.E-3
Comparison of basic EPA correlation to alternative correlations for HC

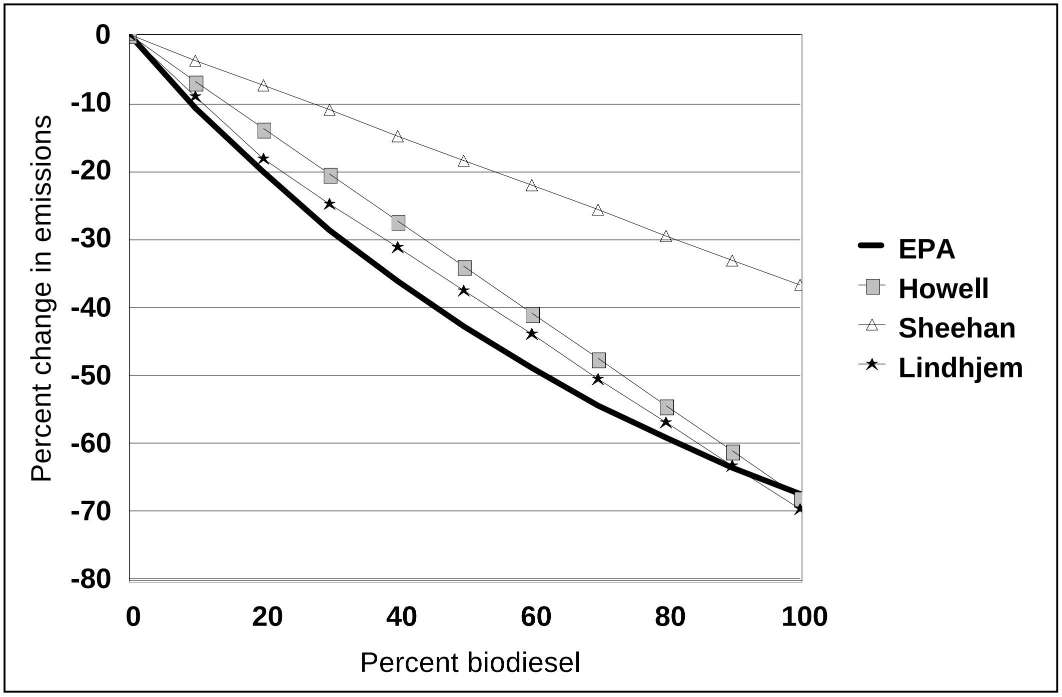
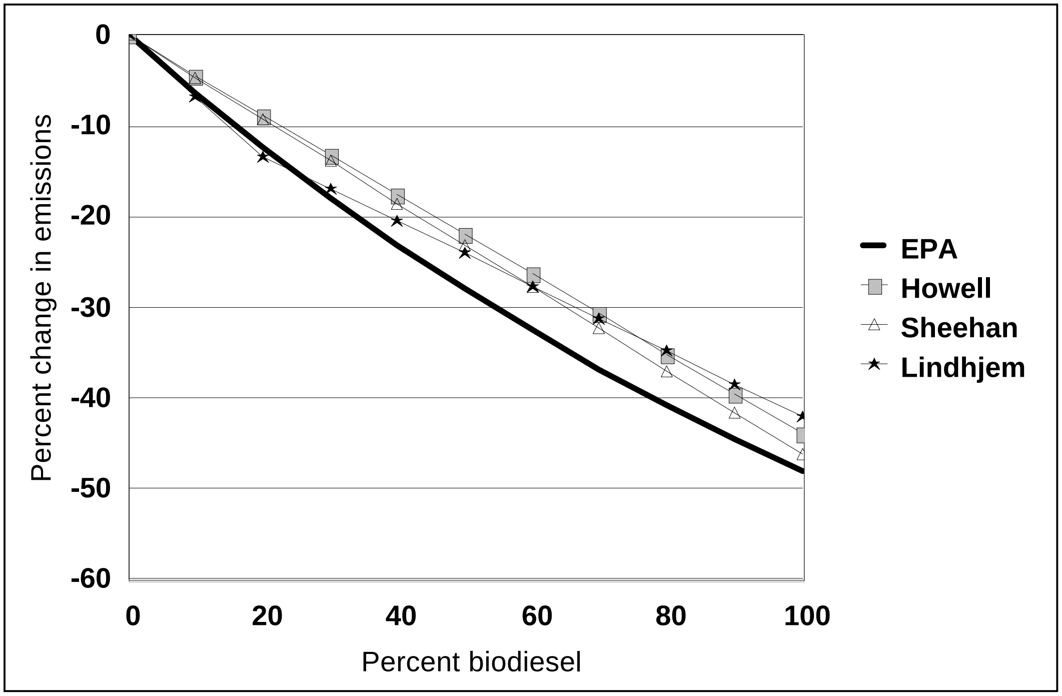


Figure IV.E-4
Comparison of basic EPA correlation to alternative correlations for CO



All four graphs show reasonable agreement between the predicted effects from our basic correlations and those from other researchers. For HC and CO, our basic correlations predict slightly larger benefits than the other correlations. This result may be due to the fact that our database contained more data, or because our analytical or statistical approach was different.

F. Applying the correlations to the in-use fleet

The composite correlations presented in Section IV.B.6 are the means through which we propose that emission impacts of biodiesel use are estimated for heavy-duty highway engines. Predictions for diesel-powered light-duty vehicles and nonroad engines are covered in Section V.

Three pollutant correlations include terms that are specific to engine standards group E (i.e. model years 1991 - 1993). As a result, these correlations are functions of the calender year for which emission benefit predictions are being made. Thus we need a means for determining what fraction of the heavy-duty highway inventory is impacted by engine standards group E. Using the inventory modeling done in the context of our rulemaking setting new standards for heavy-duty engines beginning in 2007 [66 FR 5002], we determined how the inventories were distributed among the various model years in the fleet. From this we were able to estimate the fraction of the inventory produced from either engine standards group E for any calender year. These fractions are given in Table IV.F-1.

Table IV.F-1
Yearly weighting factors for composite correlations

	Fraction of diesel highway inventory which comes from standards group E		
Pollutant	PM	CO	NOx
Factor name	k1	k2	k3
2000	0.15	0.11	0.13
2001	0.14	0.10	0.11
2002	0.13	0.10	0.10
2003	0.12	0.09	0.09
2004	0.11	0.08	0.08
2005	0.10	0.07	0.08
2006	0.10	0.06	0.07
2007	0.09	0.06	0.06
2008	0.09	0.06	0.06
2009	0.09	0.06	0.06
2010	0.09	0.06	0.05
2011	0.09	0.06	0.05
2012	0.09	0.06	0.05
2013	0.09	0.05	0.05
2014	0.09	0.05	0.05
2015	0.09	0.05	0.05
2016	0.09	0.05	0.05
2017	0.09	0.05	0.05
2018	0.09	0.05	0.05
2019	0.09	0.05	0.05
2020	0.09	0.04	0.05

The composite correlations for each pollutant can now be presented in a more useful form, as functions not only of base fuel and biodiesel source, but also of calender year. The final proposed composite correlations are shown below:

$$\begin{aligned}
&\%\ \text{change in NOx} = \\
&\quad (1 - k3) \times \{\exp[\ + 0.0010375 \times (\text{vol\% biodiesel}) \\
&\qquad + 0.0012289 \times \text{CLEAN} \times (\text{vol\% biodiesel}) \\
&\qquad - 0.0002732 \times \text{RAPE} \times (\text{vol\% biodiesel})\ ] - 1\} \times 100\% \\
&\quad + k3 \times \{\exp[\ - 0.0009795 \times \text{ANIMAL} \times (\text{vol\% biodiesel})\ ] - 1\} \times 100\%
\end{aligned}$$

$$
\begin{aligned}
\%\ \text{change in PM} = {} & \\
& (1 - k1) \times \{\exp[\ -0.0047395 \times (\text{vol\% biodiesel}) \\
& \qquad + 0.0010742 \times \text{CLEAN} \times (\text{vol\% biodiesel})\ ] - 1\} \times 100\% \\
& + k1 \times \{\exp[\ -0.0045908 \times (\text{vol\% biodiesel}) \\
& \qquad - 0.0019343 \times \text{ANIMAL} \times (\text{vol\% biodiesel})\ ] - 1\} \times 100\%
\end{aligned}
$$

$$
\begin{aligned}
\%\ \text{change in HC} = {} & \{\exp[\ -0.0118443 \times (\text{vol\% biodiesel}) \\
& \qquad + 0.0047569 \times \text{CLEAN} \times (\text{vol\% biodiesel})\ ] - 1\} \times 100\%
\end{aligned}
$$

$$
\begin{aligned}
\%\ \text{change in CO} = {} & \\
& (1 - k2) \times \{\exp[\ -0.0058238 \times (\text{vol\% biodiesel}) \\
& \qquad + 0.0010853 \times \text{CLEAN} \times (\text{vol\% biodiesel}) \\
& \qquad + 0.0017335 \times \text{RAPE} \times (\text{vol\% biodiesel})\ ] - 1\} \times 100\% \\
& + k2 \times \{\exp[\ -0.0017116 \times \text{ANIMAL} \times (\text{vol\% biodiesel})\ ] - 1\} \times 100\%
\end{aligned}
$$

where

vol% biodiesel = Value from 0 to 100
k1 = Fraction of diesel highway PM inventory which comes from engine standards group E for the calender year being considered, per Table IV.F-1
k2 = Fraction of diesel highway CO inventory which comes from engine standards group E for the calender year being considered, per Table IV.F-1
k3 = Fraction of diesel highway NOx inventory which comes from engine standards group E for the calender year being considered, per Table IV.F-1
CLEAN = 1 if the base fuel meets the conditions for "Clean" fuel given in Table III.C.2.e-1; otherwise, CLEAN = 0
ANIMAL = 1 if the biodiesel is produced from animal fat, tallow, or lard as described in Section III.C.2.c; otherwise, ANIMAL = 0
RAPE = 1 if the biodiesel is produced from rapeseed oil or canola oil, as described in Section III.C.2.c; otherwise, RAPE = 0

The categorical adjustment terms CLEAN, ANIMAL, and RAPE were chosen so that the unadjusted correlation would represent soybean-based biodiesel added to an "average" base fuel. Since soybeans are currently the largest single feedstock for biodiesel in the U.S.[f] and the majority of in-use fuels used for blending with biodiesel can be categorized as "average," the unadjusted correlation is the best representation of biodiesel effects on emissions when the actual biodiesel feedstock and base fuel properties are unknown. This case is shown in Table IV.F-2 for the year 2003 for a 20 vol% biodiesel blend.

---

[f] According to the National Biodiesel Board, 80% of current biodiesel is derived from soybeans, 19% from yellow grease, and the remaining 1% from other feedstocks.

Table IV.F-2
Default emission effects of 20 vol% biodiesel in 2003
(assumes soybean-based biodiesel added to an "average" base fuel)

	Percent change in emissions
NOx	+ 2.0 %
PM	- 10.1 %
HC	- 21.1 %
CO	-11.0 %

Section V: Biodiesel Effects on Light-Duty Vehicles and Nonroad

Because the amount of heavy-duty highway engine data in our database far surpassed the amount of data for other equipment types, we opted to base our primary analyses on heavy-duty highway engine data only. However, nonroad diesel engines and, to a much lesser extent, light-duty diesel vehicles, also contribute to the total mobile source inventories for regulated pollutants. We therefore compared the emission impacts of biodiesel use estimated for heavy-duty highway engines to data collected on nonroad engines and light-duty vehicles. Based on the analyses described in this Section, we have concluded that there is insufficient support for extrapolating heavy-duty highway engine effects to nonroad or light-duty.

A. Methodology

We first converted all nonroad and light-duty vehicle emissions data from absolute g/bhp-hr or g/mile to percent change in emissions. This process began by identifying the base fuel to which biodiesel had been added for each vehicle/test cycle combination, and averaging any repeat measurements made on this base fuel. The percent change values for every emission measurement were then calculated with respect to this average base fuel emissions measurement, producing "observed" values. To calculate "predicted" values, we used our composite correlations to estimate the percent change in emissions of each pollutant for the specific biodiesel concentration associated with each nonroad or light-duty vehicle test.

We compared the predicted and observed percent change emission values in several different ways to determine if our heavy-duty highway engine-based correlations could be reasonably applied to either nonroad engines or light-duty vehicles. The different types of comparisons are listed in Table V.A-1.

Table V.A-1
Methods for comparing nonroad and light-duty data to the composite correlations

Comparison	Purpose
Graphical	Provided a qualitative assessment of the match between predicted and observed values, as well as scatter in the data and potential bias.
Paired t-tests of predicted versus observed percent change values	Provided a more quantitative assessment of potential bias between the composite model and the data
Ratio of mean absolute value residuals	Provided a quantitative assessment of scatter

The ratio of residuals was calculated from the following formula:

$$Ratio = \frac{\sum_{i=1}^{n_d} \left|R_{d_i}\right| / n_d}{\sum_{j=1}^{n_m} \left|R_{m_j}\right| / n_m}$$

where:

Rd_i = Residual for the nonroad or light-duty data being evaluated, for observation i
= Observed percent change for nonroad or light-duty observation i minus composite correlation's predicted percent change associated with observation i
$|Rd_i|$ = Absolute value of Rd_i
Rm_j = Residual for the heavy-duty highway engine data on which the composite correlation was based, for observation j
= Observed percent change for highway engine observation j minus composite correlation predicted percent change for observation j
$|Rm_j|$ = Absolute value of Rm_j
n_d = Total number of biodiesel observations for the nonroad or light-duty data being evaluated
n_m = Total number of biodiesel observations for the heavy-duty highway engine data on which the composite correlation was based

A ratio significantly higher than 1.0 indicated that the scatter in the nonroad and/or light-duty vehicle data was significantly greater than that for the data on which the composite correlations were based. Such a result would call into question whether the estimated emission effects for heavy-duty highway engines could appropriately be applied to nonroad or light-duty. For example, a ratio of 1.5 would suggest that the scatter in the nonroad or light-duty data about the composite correlation was 50% greater than the scatter in the heavy-duty highway engine data on which that composite correlation was based. A ratio of 2.0 would suggest that the scatter in the nonroad or light-duty data about the composite correlation was twice that of the heavy-duty highway engine data. For our purposes we considered a ratio of 1.5 to be noteworthy and 2.0 to be substantial.

The primary purpose in conducting this analysis was to determine if the composite correlations presented in Section IV.F for the heavy-duty highway fleet can be used to predict emission impacts of biodiesel for nonroad engines or light-duty highway vehicles. However, the composite correlations include multiplicative factors k1 - k3 that complicate the analysis. These factors account for the idiosyncratic effects of biodiesel on heavy-duty highway engines having model years 1991 through 1993. This group of model years was identified as engine standards group E during our derivation of the composite correlations, and represents one particular set of heavy-duty highway engine certification standards as shown in Table II.C-1. However, these emission standards are not directly relevant for nonroad engines or light-duty vehicles. Thus the values for multiplicative factors k1 - k3 as presented in Table IV.F-1 cannot be applied to nonroad or light-duty vehicles. Among the various possible approaches for addressing this issue, the most straightforward is to use the composite correlations presented in Section IV.B.6 instead of those in Section IV.F (which were modified to more directly apply to the in-use fleet of heavy-duty highway engines), and to set the variable "GROUP E" equal to zero. This approach recognizes the fact that there is no group of model years among nonroad engines or light-duty highway vehicles that is uniquely comparable to highway engine standards group E.

B. Effects of biodiesel on nonroad engines

Most large nonroad diesel engines use technologies similar to those found in heavy-duty highway diesel engines, although in a given year, the highway engine technology is generally more advanced. On this basis we might expect biodiesel to produce similar emission effects for heavy-duty nonroad and heavy-duty highway engines. Additionally, in both this work and previous analyses, we have found few distinctions between different highway engine technologies in terms of fuel effects on emissions (i.e. very few engine standard group terms were significant, as shown in Table IV.B.3-1). Given this result among different highway engine technologies, one might expect that the distinctions between highway engine technology and nonroad technology would similarly not be important in the context of evaluating the effects of biodiesel on emissions. The analyses we conducted for heavy-duty nonroad engines were intended to determine if this expectation was reasonable.

Our database contained only 8 biodiesel observations for heavy-duty nonroad engines, which falls below our minimum data criteria. Thus it might not be reasonable to expect the biodiesel effects exhibited by these nonroad engines to compare favorably with effects predicted by the composite correlations, simply due to variability in emission measurements. In addition, the nonroad engines were all tested on the 8-mode steady-state cycle used for nonroad certification, whereas the heavy-duty highway engines were tested primarily on the FTP (and, to a lesser extent, the European 13-mode steady-state test). The difference in test cycles may also contribute to differences between predicted and observed values for nonroad engines. This is especially true for PM and CO, which we determined in Section IV.B.1 exhibited different effects of biodiesel on emissions for transient versus steady-state operation.

We first developed graphical comparisons between our modeled effects for heavy-duty highway engines and the nonroad data. The results are shown in Figures V.B-1 through V.B-4.

Figure V.B-1
Predicted vs observed NOx for nonroad

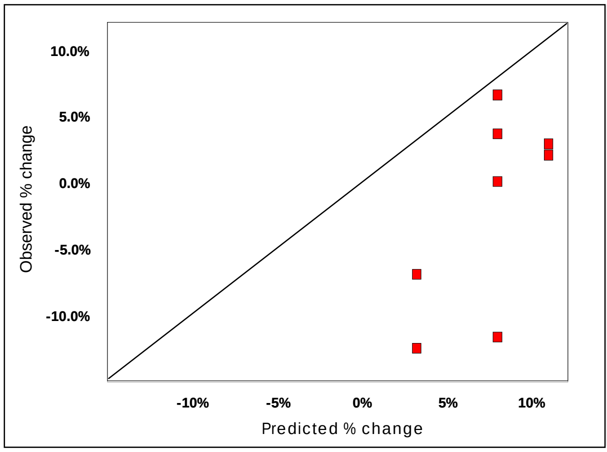


Figure V.B-2
Predicted vs. observed PM for nonroad

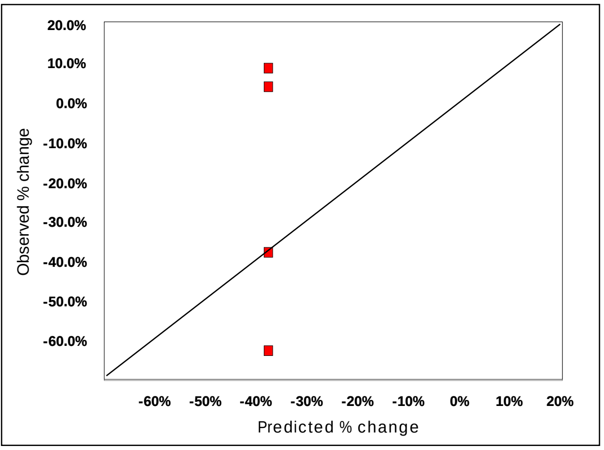


Figure V.B-3
Predicted vs observed CO for nonroad

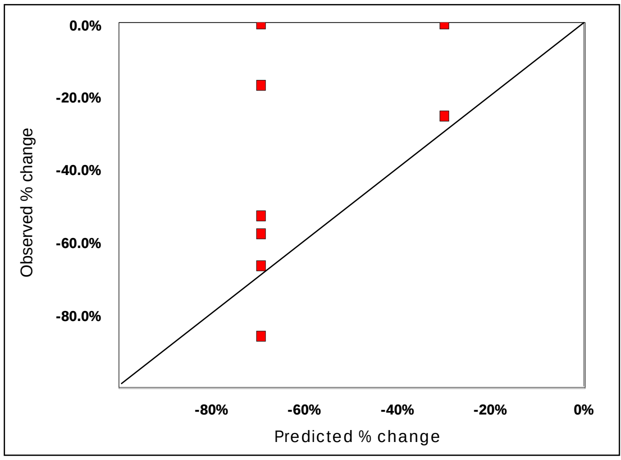
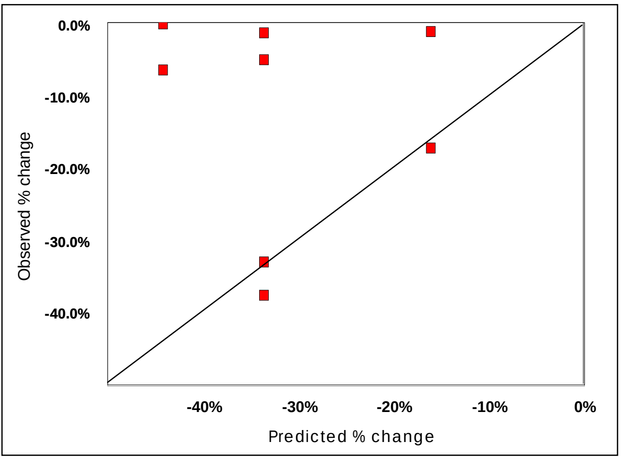


Figure V.B-4
Predicted vs. observed HC for nonroad



In order to conclude that nonroad effects of biodiesel use could be reasonably predicted by the composite correlations based on heavy-duty highway engines, the data should exhibit equal scatter about the diagonal line. In the figures above, it appears that there is bias between the predicted and observed values. To verify that this bias is significant, we used paired t-tests to compare predicted and observed percent change values. The results are shown in Table V.B-1.

Table V.B-1
Paired t-tests for nonroad data

	Probability that predicted and observed values are different
NOx	0.997
PM	0.577
CO	0.975
HC	0.934

The values in Table V.B-1 confirm that the composite correlations do a poor job of representing the nonroad engine effects in our database. PM may appear to be an exception. However, there are only four observations for PM, raising the possibility that the overlap between predicted and observed values in Figure V.B.2 may be a statistical anomaly. The overlap is also difficult to interpret given the established differences between transient and steady-state results for PM as described in Section IV.B.1.

We also examined the ratio of mean absolute residuals to determine if the scatter in the nonroad data about the composite correlations was comparable to that exhibited by the highway data on which those composite correlations were based. Table V.B-2 provides these ratios for the nonroad data in our database.

Table V.B-2
Ratio of mean absolute value residuals for nonroad

NOx	2.60
PM	5.23
CO	2.51
HC	1.55

Since these ratios are significantly higher than 1.0, we conclude that the scatter in the nonroad data around the composite correlations is substantial. This result might be expected given the high degree of bias exhibited in Figures V.B-1 through V.B-4, since that bias contributes to this particular measure of scatter. Therefore, we also calculated these ratios using an alternative nonroad-only set of correlations. In this analysis, we used a simple linear least-squares regression to produce correlations based on only the nonroad data. Although the available nonroad data is too limited to produce correlations that could be used to represent the entire nonroad fleet, these correlations do allow the bias exhibited in Figures V.B-1 through V.B-4 to be eliminated. The ratios are then the mean absolute value residuals for the nonroad data around the nonroad correlations, compared to the mean absolute value residuals for the highway data around the highway composite correlations. The result is a set of ratios that might be considered a more appropriate measure of relative scatter. The results are shown in Table V.B-3.

Table V.B-3
Ratio of mean absolute value residuals for nonroad
using least-squares regression for nonroad data

NOx	1.64
PM	3.63
CO	1.63
HC	1.25

The values in Table V.B-3 still indicate a significantly higher amount of scatter in the nonroad data relative to scatter in the highway data. Given the limited number of nonroad observations and the higher level of scatter exhibited by that nonroad data, it would be reasonable to conclude that there is insufficient information to determine whether the composite correlations based on highway engines can be used to represent nonroad engine effects of biodiesel use.

We also investigated whether the bias exhibited in Figures V.B-1 through V.B-4 might be a result of differences in test cycles. As stated above, the nonroad data was collected on an 8-mode steady-state nonroad cycle, whereas most of the heavy-duty highway data was collected on the transient FTP. Since there was some highway engine data collected on the European 13-mode steady-state cycle ("R49"), we regenerated the basic correlations described in Section IV.A, but using only the highway engine R49 data. The results are shown in Table V.B-4.

Table V.B-4
Coefficients for basic emission correlations using R49 data
% change = {exp[a × (vol% biodiesel)] - 1} × 100%

	Coefficient 'a'	P-value
NOx	0.0014502	0.0001
PM	- 0.001108	0.2990
CO	- 0.00122	0.4934
HC	- 0.004832	0.0402

The coefficients in Table V.B-4 are directionally consistent with those in Table IV.A-1. The fact that the PM and CO correlations are not significant is consistent with our investigation of test cycle effects as described in Section IV.B.1.

We used these R49-based highway engine correlations to compare predicted emission impacts to those observed for nonroad engines. The results are shown in Figures V.B-5, V.B-6, V.B-7, and V.B-8. Note that these graphical comparisons use the actual estimated coefficients listed in Table V.B-4, rather than assigning a zero coefficient to the PM and CO correlations. The corresponding paired t-tests and the ratio of mean absolute value residuals are given in Table V.B-5.

Figure V.B-5
Predicted vs observed NOx for nonroad using R49 correlation

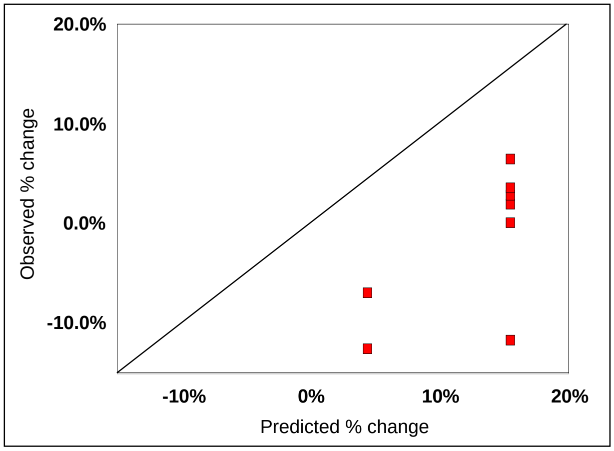


Figure V.B-6
Predicted vs observed PM for nonroad using R49 correlation

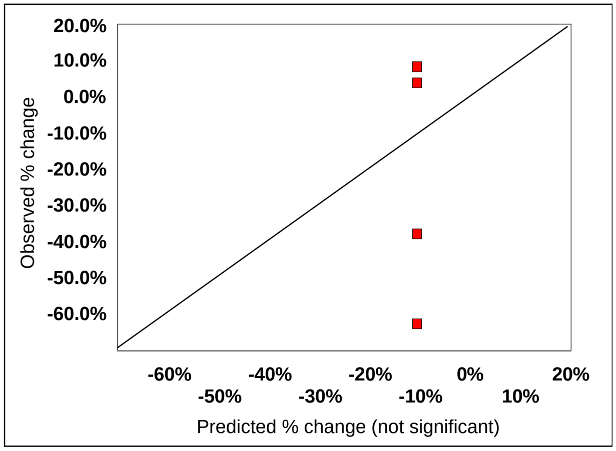


Figure V.B-7
Predicted vs observed CO for nonroad using R49 correlation

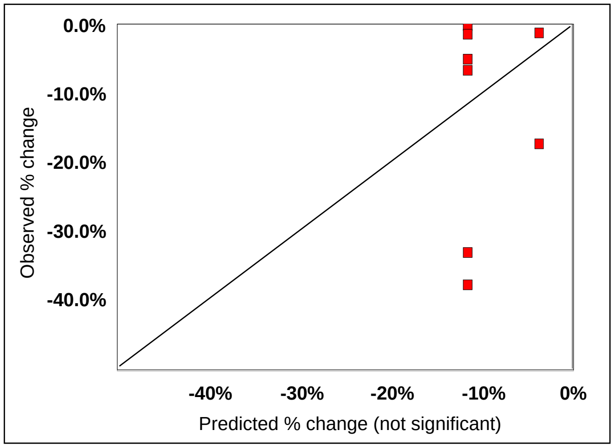


Figure V.B-8
Predicted vs observed HC for nonroad using R49 correlation

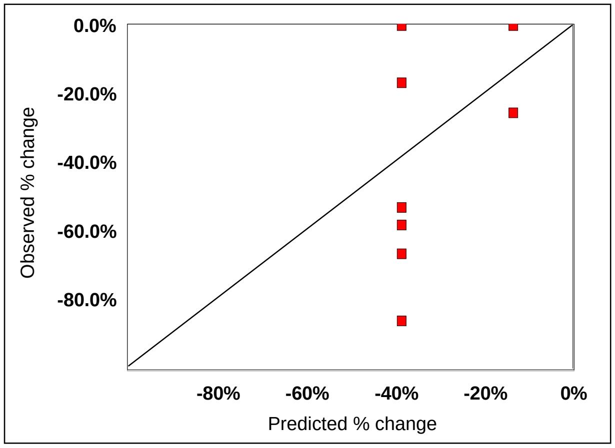


Table V.B-5
Evaluation of nonroad data using R49 correlation

	Paired t-tests: Probability that predicted and observed values are different		Ratio of mean residuals
	Ignoring P-values	Zero effect for not significant correlations	
NOx	0.999	0.999	4.06
PM	0.447	0.709	3.73
CO	0.425	0.951	1.49
HC	0.426	0.426	1.47

The paired t-tests indicate that the predicted and observed NOx values are clearly different from one another. The best-fit CO correlation with a zero coefficient also produces a statistically significant difference between predicted and observed values. Since this analysis was intended to remove test cycle effects from the comparison of nonroad and highway responses to biodiesel, we conclude that the bias exhibited in Figures V.B-1 and V.B-3 for NOx and CO, respectively, is most likely due to factors other than test cycle differences, i.e. the results appear to be due to differences in the way that nonroad and highway engines respond to biodiesel. However, given the high amount of scatter in the nonroad data relative to highway data, it is not clear that the true cause of the bias can be discerned from the data in our database. As a result, we cannot draw a clear conclusion regarding the degree to which highway and nonroad engines respond similarly to biodiesel, and there is some evidence that they respond differently. On the basis of these analyses, then, we cannot say with confidence that the composite correlations presented in Section IV.B.6 can provide accurate predictions of biodiesel effects on emissions for nonroad engines.

There is an alternative analytical approach that could be used to produce a single set of correlations representing both highway and nonroad engines. In this approach, the nonroad data would be pooled with the highway data when the SAS procedure proc mix is run. Adjustment terms for nonroad could then be permitted in these regression analyses. If none of the nonroad adjustment terms were significant, then we might conclude that nonroad effects of biodiesel cannot be distinguished from heavy-duty highway effects of biodiesel, and that the composite correlations presented in Section IV.B.6 can be applied to heavy-duty nonroad engines. Such an approach would be more appropriate for NOx and HC which appear to exhibit similar responses to biodiesel for transient and steady-state cycles. It may be less appropriate to use this approach for PM and CO. Regardless, given that the amount of nonroad data is substantially less than that for highway, it is unclear if this approach would yield correlations (regardless of whether adjustment terms for nonroad are significant) that would be representative of the nonroad fleet.

Finally, the preceding analyses applied only to heavy-duty nonroad engines. Light-duty engines are sufficiently different from heavy-duty that we would not expect the effects of biodiesel on emissions to be the same for heavy-duty highway and light-duty nonroad. In this case, we would propose that biodiesel emission effects estimated for heavy-duty highway engines not be applicable to light-duty nonroad engines unless strong evidence to the contrary exists. Given that our database contained no light-duty nonroad engines, no evaluation was possible, and we therefore propose that our models not be applied to light-duty nonroad engines.

C. Effects of biodiesel on light-duty highway vehicles

As for light-duty nonroad engines, light-duty highway diesel engines are sufficiently different from heavy-duty that we would not expect the effects of biodiesel on emissions to be the same for heavy-duty highway and light-duty highway. Thus we would propose that biodiesel emission effects estimated for heavy-duty highway engines not be applicable to light-duty highway engines unless strong evidence to the contrary exists.

Our database contained only three biodiesel observations for light-duty vehicles. All three observations were collected on the same biodiesel blend, engine, and test cycle. Given the extreme paucity of light-duty data, we do not believe that a conclusive determination can be reached regarding the appropriateness of using the highway composite correlations given in Section IV.B.6 to represent biodiesel effects on light-duty vehicles. We therefore default to our recommendation that the composite correlations not be used to represent light-duty vehicles.

However, in the interest of completeness, we present a comparison of observed and predicted % change values for the three light-duty vehicle observations in our database. The comparisons are shown graphically in Figures V.C-1 through V.C-4. Table V.C-1 presents the accompanying t-tests and ratio of mean residuals.

Figure V.C-1
Predicted vs observed NOx for light-duty

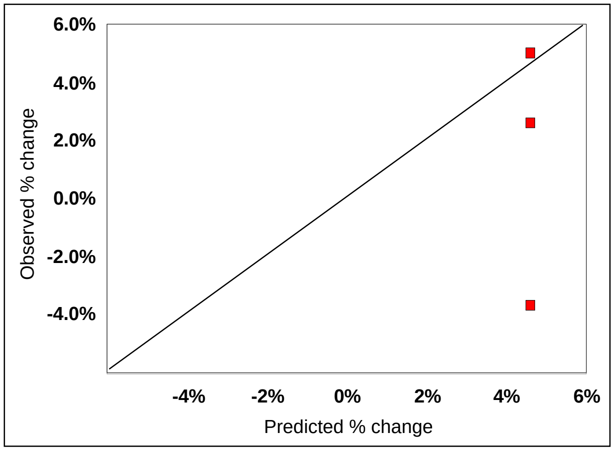


Figure V.C-2
Predicted vs observed PM for light-duty

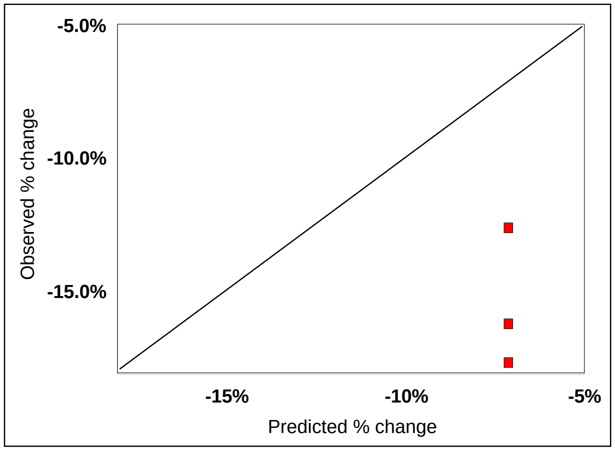


Figure V.C-3
Predicted vs observed for CO

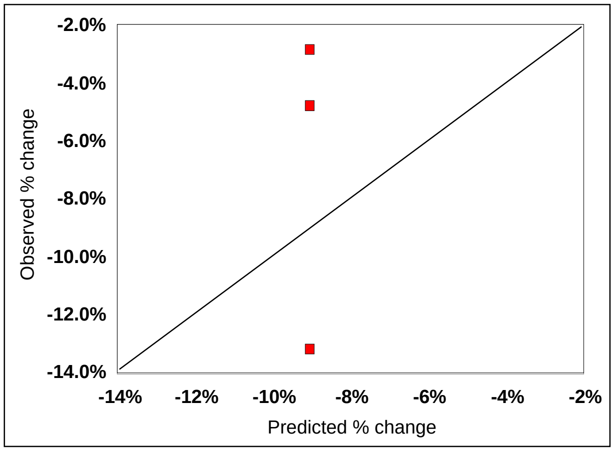


Figure V.C-4
Predicted vs observed for HC

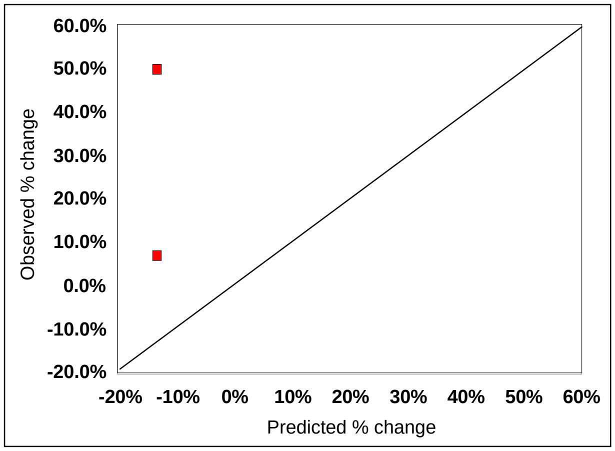


Table V.C-1
Evaluation of light-duty data

	Paired t-tests: Probability that predicted and observed values are different	Ratio of mean residuals
NOx	0.664	0.98
PM	0.969	1.10
CO	0.428	0.60
HC	0.924	2.97

Section VI: Biodiesel Effects On Gaseous Toxics

In addition to impacts on regulated pollutants, we also investigated the impact that biodiesel has on emissions of unregulated hazardous air pollutants, hereafter referred to as toxics. Most of the studies in our database focused only on the effects of regulated pollutants. The total amount of data on toxics effects was much smaller than that for regulated pollutants, and as a result we took a different approach to evaluating biodiesel effects on toxics. We treat our conclusions regarding the effects of biodiesel on toxics as preliminary and only potentially indicative of the true effects, due to the limited nature of the data.

A. Toxic Pollutants Evaluated

Hydrocarbons include many different individual toxic compounds. For the purpose of evaluating biodiesel effects on toxics, we focused on mobile source air toxics (MSATs) as defined in a recent rulemaking[8]. MSATs are significant contributors to toxic emission inventories, and are known or suspected to cause cancer or other serious health effects. There are 21 MSATs as shown in Table VI.A-1.

Table VI.A-1
List of 21 Mobile Source Air Toxics (MSATs)

- Acetaldehyde
- Acrolein
- Arsenic Compounds[1]
- Benzene
- 1,3-Butadiene
- Chromium Compounds[1]
- Diesel Particulate Matter + Diesel Exhaust Organic Gases (DPM + DEOG)
- Dioxin/Furans[2]
- Ethylbenzene
- Formaldehyde

[1] Although the different metal compounds generally differ in their toxicity, the onroad mobile source inventory contains emissions estimates for total metal compounds (i.e., the sum of all forms).

[2] This entry refers to two large groups of chlorinated compounds. In assessing their cancer risks, their quantitative potencies are usually derived from that of the most toxic, 2,3,7,8-tetrachlorodibenzodioxin.

- n-Hexane
- Lead Compounds[1]
- Manganese Compounds[1]
- Mercury Compounds[1]
- MTBE
- Naphthalene
- Nickel Compounds[1]
- POM[3]
- Styrene
- Toluene
- Xylene

[3] Polycyclic Organic Matter includes organic compounds with more than one benzene ring, and which have a boiling point greater than or equal to 100 degrees centigrade. A group of seven polynuclear aromatic hydrocarbons, which have been identified by EPA as probable human carcinogens, (benz(a)anthracene, benzo(b)fluoranthene, benzo(k)fluoranthene, benzo(a)pyrene, chrysene, 7,12-dimethylbenz(a)anthracene, and indeno(1,2,3-cd)pyrene) are used here as surrogates for the larger group of POM compounds.

We reviewed the available biodiesel studies to determine which studies had actually measured and reported any of the MSATs listed in Table VI.A-1. To be considered in this analysis, the studies also had to conform to the criteria for choosing data sources, as described in Section II.A. The resulting list of seven studies used for this toxics analysis is given in Appendix D.

Of the 21 MSATs, six are metals. Since none of the studies we reviewed contained measurements of metal emissions, we cannot directly determine the relationship between biodiesel use and metal emissions. However, biodiesel is essentially metal-free, so we would expect reduced emissions of metals for biodiesel blends, if indeed conventional diesel fuel or any additives used in conventional diesel contains metals. Similarly, emissions of MTBE are not an issue for diesel fuels since MTBE is only used as an additive for gasoline. Of the remaining fourteen MSATs, we have emission measurements for eleven. Table VI.A-2 lists the eleven MSATs that we therefore investigated for biodiesel effects.

Table VI.A-2
MSATs Investigated for Biodiesel Effects

acetaldehyde acrolein benzene 1,3-butadiene ethylbenzene formaldehyde	n-hexane naphthalene styrene toluene xylene

B. Analytical Approach

The available toxics data was collected on different engines, test cycles, and base fuels. For our purposes, we focused on those occasions in which a biodiesel blend was formulated using the same base fuel to which the blend was being compared. We then summarized the data in a frequency table to determine how many toxic measurements were available. Table VI.B-1 lists the number of times a toxic compound was measured at specific biodiesel concentrations. Every observation was a unique measurement of a given fuel on a given engine; there were no repeat emission measurements.

Table VI.B-1
Number of toxic observations

	Percent biodiesel				Total
	0 % (base)	20%	50%	100%	
Acetaldehyde	15	15	0	7	37
Acrolein	10	10	0	7	27
Benzene	9	9	0	7	25
1,3-Butadiene	7	7	0	5	19
Ethylbenzene	7	7	0	5	19
Formaldehyde	15	15	0	7	37
n-Hexane	9	9	0	7	25
Naphthalene	11	9	4	9	33
Styrene	4	4	0	2	10
Toluene	9	9	0	7	25
Xylene	9	9	0	7	25
Total	105	103	4	70	282

Before proceeding with an evaluation of each toxic separately, we first investigated the degree to which the aggregated sum of all toxics were correlated with biodiesel concentration. Although individual toxics vary from one another in terms of health effects and there is no reason to believe that one toxic compound will respond to the presence of biodiesel identically to another toxic compound, an analysis of the aggregated sum of toxics was appropriate for several reasons. First, the amount of available data for total toxics was much larger than that for any individual toxic compound, so that correlations are more likely to be statistically significant. As a result, a correlation between biodiesel concentration and total toxics provides a means for judging the overall toxics impacts of biodiesel without making statements regarding how individual toxic compounds are affected by the presence of biodiesel. We recognize that individual toxics may increase or decrease when biodiesel is blended with diesel fuel, and of those that decrease, the magnitude of that decrease will vary from one toxic to another. Thus the correlation of biodiesel concentration with total toxics is bound to be smaller in magnitude than some individual toxics, and larger than others.

Second, since the toxics in Table VI.B-1 are components of total hydrocarbon emissions, this approach allowed us to compare biodiesel effects on total toxics to biodiesel effects on total hydrocarbons in order to determine if the combustion mechanisms could be said to differ between the two (overlapping) groups. This latter issue is important because, in the absence of sufficient or conclusive information on toxics emissions from diesel engines, we often make the assumption that the ratio of toxics to total hydrocarbons remains constant.

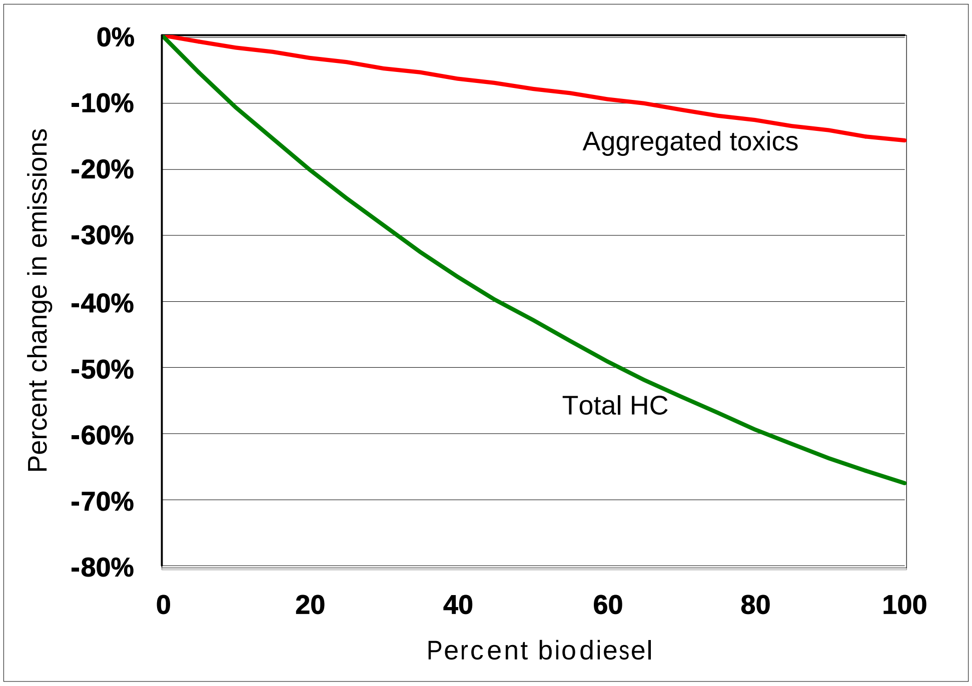
There were no repeat measurements among the toxics data we evaluated. Therefore, there was no need to average repeat base fuel emission measurements. There did not appear to be sufficient information to directly evaluate such issues as test cycle effects, engine technology, base fuel, or type of biodiesel. However, these variables could still have an impact on the response of toxics to biodiesel. In order to control for these variables as well as individual engine effects, we first calculated the percentage change in emissions of each toxic compound for every available toxic emissions measurement at a given non-zero biodiesel concentration, with respect to the toxics measurements for the base fuel. This simplifying approach was deemed appropriate for this preliminary analysis of biodiesel effects on toxics, since the limited amount of toxics data may not be sufficient to draw specific, quantified conclusions applicable to the entire in-use fleet. Rather, our analysis was intended to investigate directional effects and the potential magnitude of those effects, with the understanding that a more comprehensive toxics database may be necessary to produce more robust estimates.

We identified and eliminated outliers that exceeded four standard deviations from the mean (two observations were eliminated thus), and then used linear least-squares regression to correlate percent change in aggregated (total) toxics with biodiesel concentration. The resulting correlation was statistically significant, and is shown below:

$$\%\ \text{change in total toxics} = -0.001580 \times (\%\ \text{biodiesel})$$

This correlation was compared to the basic model for total hydrocarbons given in Section IV.A. The result is shown in Figure VI.B-1.

Figure VI.B-1
Comparison of aggregated toxics and total HC



Based on this analysis, we could reasonably conclude that total toxics are reduced when biodiesel is added to conventional diesel fuel. This conclusion differs from the outcome that would result if one assumed that the mass ratio of toxics to hydrocarbons was constant, i.e. independent of biodiesel concentration. On the contrary, our analysis suggests that the mass ratio of total toxics to total hydrocarbons actually increases with the addition of biodiesel, though apparently not to the point of causing total toxics emissions to increase over baseline levels.

The above analysis of total toxics gives no indication of how individual toxics might respond to the addition of biodiesel to conventional diesel fuel. In fact, it is possible that emissions of one toxic compound might increase with increasing biodiesel concentration even though the aggregate of all toxics decreases. It therefore seemed prudent to investigate individual toxics despite the relative paucity of data in comparison to total toxics and total hydrocarbons.

We took several different approaches to evaluating data for individual toxics because the data was so limited. Consideration of the results of all of these analytical approaches was used in drawing conclusions about the effect of biodiesel on individual toxics. These approaches are listed below:

a) Correlation of mass ratio of toxic/HC with biodiesel concentration
b) Correlation of % change in toxics emissions with biodiesel concentration
c) Binomial analysis of increases and decreases in toxics for biodiesel blends
d) Difference in average toxics effects at 20% biodiesel and 100% biodiesel

The results of these analyses are shown in Tables VI.B-2, VI.B-3, VI.B-5, and VI.B-6. A discussion of how the results were together considered to draw conclusions about individual toxics is given in Section VI.C. Note that '% biodiesel' refers to values between 0 and 100.

Table VI.B-2
Correlation of toxic/HC ratios with % biodiesel
toxic/HC ratio = a × (% biodiesel) + b

	a	b	P-value for a
Acetaldehyde	0.000403	0.051071	0.03
Acrolein	0.000142	0.026455	0.06
Benzene	0.000097	0.012679	0.01
1,3-Butadiene	0.000015	0.014188	0.33
Ethylbenzene	-0.000032	0.003687	0.02
Formaldehyde	0.000917	0.129502	0.02
n-Hexane	0.000011	0.000673	0.05
Naphthalene	0.000011	0.001104	0.01
Styrene	0.000021	0.000694	0.02
Toluene	0.000022	0.009908	0.26
Xylene	-0.000054	0.006798	0.02

Table VI.B-3
Correlation of % change in toxics with % biodiesel
% change in toxics = c × (% biodiesel) × 100%

	c	P-value for c
Acetaldehyde	-0.001606	0.05
Acrolein	-0.000846	0.21
Benzene	0.000390	0.40
1,3-Butadiene	-0.000132	0.48
Ethylbenzene	-0.006970	0.00
Formaldehyde	-0.001696	0.00
n-Hexane	-0.002381	0.12
Naphthalene	-0.002847	0.04
Styrene	0.003501	0.05
Toluene	0.001750	0.19
Xylene	-0.004078	0.01

Note that most of the values for 'c' in Table VI.B-3 are negative, indicating that emissions of the toxic compound are generally reduced when biodiesel is added to conventional diesel fuel.

Of the three positive values, two were not statistically significant. This result is consistent with the observation that total toxics appear to be reduced with increasing biodiesel concentration, as shown in Figure VI.B-1.

The correlations in Table VI.B-2 can also be used in combination with the basic HC correlation from Section IV.A.1 to estimate the impact that biodiesel has on toxics emissions. This approach yields a means for checking that the sign of the coefficients in Table VI.B-3 make sense. To do this, we first assumed a basic HC emission rate of 1.0 g/bhp-hr (although the specific value has no impact on the calculation). The basic HC correlation from Table IV.A.1-1 was then converted into one giving emissions of HC in g/bhp-hr as follows:

$$\text{HC emissions (g/bhp-hr)} = 1 + 1 \times \{\exp[-0.011195 \times (\text{vol\% biodiesel})] - 1\}$$

This equation was then combined with each of the correlations in Table VI.B-2 to yield an equation giving emissions of each toxic compound in g/bhp-hr as a function of vol% biodiesel:

$$\text{Toxic (g/bhp-hr)} = \\ \big\{a \times (\text{\% biodiesel}) + b\big\} \times \big\{1 + 1 \times \{\exp[-0.011195 \times (\text{vol\% biodiesel})] - 1\}\big\}$$

Finally, this equation was converted from g/bhp-hr to percent change. The resulting equation can be rearranged to the following form:

$$\text{\% change in toxics} = \\ \big\{\{a/b \times (\text{\% biodiesel}) + 1\} \times \exp[-0.011195 \times (\text{vol\% biodiesel})] - 1\big\} \times 100\%$$

where the coefficients 'a' and 'b' are drawn from Table VI.B-2. We were able to use this equation to estimate directional effects for each toxic (i.e. increase or decrease when biodiesel is added to diesel fuel), and then to compare these directional effects with the coefficients for 'c' in Table VI.B-3 for consistency. The results are shown in Table VI.B-4, where "Table VI.B-2 correlations" refers to the equation above. Note that we did not regard statistical significance of the regressions for this comparison.

Table VI.B-4
Comparison of directional effects for toxics using two different approaches

	Toxics emissions effect with increasing biodiesel for...	
	Table VI.B-2 correlations	Table VI.B-3 correlations
Acetaldehyde	Decrease	Decrease
Acrolein	Decrease	Decrease
Benzene	**Decrease**	**Increase**
1,3-Butadiene	Decrease	Decrease
Ethylbenzene	Decrease	Decrease
Formaldehyde	Decrease	Decrease
n-Hexane	Decrease	Decrease
Naphthalene	Decrease	Decrease
Styrene	Increase	Increase
Toluene	**Decrease**	**Increase**
Xylene	Decrease	Decrease

The comparison in Table VI.B-4 indicates an inconsistency for benzene and toluene. The implications are discussed in more detail in Section VI.C.3 below.

Table VI.B-5
Binomial distributions of toxic increases versus decreases

	Number of biodiesel observations of % change from baseline which are...		Probability that a random distribution would produce this result
	Positive	Negative	
Acetaldehyde	7	15	0.13
Acrolein	7	10	0.63
Benzene	7	5	0.77
1,3-Butadiene	7	5	0.77
Ethylbenzene	2	10	0.04
Formaldehyde	8	14	0.29
n-Hexane	3	8	0.23
Naphthalene	6	12	0.24
Styrene	2	2	n/a
Toluene	8	8	n/a
Xylene	2	10	0.04

The probabilities in Table VI.B-5 give an indication of how likely it is that the direction (not the magnitude) of the true effect can be discerned from the direction of the individual observations. Normally a value of 0.10 (corresponding to a 90% confidence interval) would indicate a statistically significant result. Given the small amount of data and the notorious variability in toxics measurements, we have used the binomial distribution results in a less rigorous fashion as a check on the correlations given in Table VI.B-3. This comparison is discussed for each individual toxic in Section VI.C below.

Table VI.B-6
Difference in average toxics effects at two biodiesel blend levels

	Average % change compared to base fuel		Consistent trend?
	20% biodiesel	100% biodiesel	
Acetaldehyde	-7.1 %	-14.4 %	Yes
Acrolein	-1.5 %	-8.5 %	Yes
Benzene	16.5 %	-0.8 %	No
1,3-Butadiene	39.0 %	-12.3 %	No
Ethylbenzene	-44.9 %	-61.0 %	Yes
Formaldehyde	-7.8 %	-15.1 %	Yes
n-Hexane	-48.7 %	-12.1 %	?
Naphthalene	-13.8 %	-26.7 %	Yes
Styrene	-3.7 %	39.3 %	No
Toluene	19.9 %	13.3 %	?
Xylene	-12.3 %	-39.5 %	Yes

For n-hexane and toluene in Table VI.B-6, the directional effects are consistent between the 20% and 100% biodiesel blends, but the magnitude of those effects are not consistent with expectations (we would expect the magnitude of the effect at 100% biodiesel to be larger than that at 20% biodiesel). Such results could easily be an outgrowth of the paucity of data. Regardless, the trend for these two toxics is not wholly inconsistent, a fact taken into account in our discussions in the next Section.

C. Conclusions for individual toxics

Since the amount of available data on toxic impacts of biodiesel use is so limited, we treat our conclusions regarding the effects of biodiesel on toxics as preliminary and only potentially indicative of the true effects. We have some confidence that the impact of biodiesel on total toxics is beneficial, i.e. total toxics are on average reduced when biodiesel is added to conventional diesel fuel. We have less confidence in our conclusions for individual toxic compounds. It is our hope that additional toxics data will be generated by independent researchers in the near future, allowing a more robust and definitive analysis of biodiesel effects on toxics.

Because the available toxics data was so limited, we determined that a tiered approach to drawing conclusions about biodiesel effects on toxics would be appropriate. We have defined three tiers corresponding to three different qualitative levels of confidence in the conclusions. They are described below:

> Tier 1: Includes those toxic compounds for which the analyses described in Section VI.B above appear to be largely consistent with one another. The effect of biodiesel on toxics can be quantified with reasonable confidence.
>
> Tier 2: Includes those toxic compounds for which the analyses may not be entirely consistent with one another, or not statistically significant. However, the effect of biodiesel on toxics can still be estimated qualitatively in terms of directional effects.
>
> Tier 3: Includes those toxic compounds for which the analyses are significantly in conflict with one another. No clear conclusions can be drawn.

The eleven toxic compounds evaluated in this analysis were assigned to one of the three tiers, as shown in Table VI.C-1. Each of the following three subsections describes our conclusions for the three tiers of toxics.

Table VI.C-1
Tier assignments for toxics

Acetaldehyde	Tier 1
Acrolein	Tier 2
Benzene	Tier 3
1,3-Butadiene	Tier 3
Ethylbenzene	Tier 1
Formaldehyde	Tier 1
n-Hexane	Tier 2
Naphthalene	Tier 1
Styrene	Tier 2
Toluene	Tier 3
Xylene	Tier 1

1. Tier 1 toxics

Tier 1 toxics include those toxic compounds for which the analyses described in Section VI.B above appear to be largely consistent with one another. The effect of biodiesel on toxics can be quantified with reasonable confidence. The toxics which we have determined fall into this tier include acetaldehyde, ethylbenzene, formaldehyde, naphthalene, and xylene. Table VI.C.1-1 repeats the correlations presented in Table VI.B-3 for these five toxics. Note that '%

biodiesel' is given as a value between 0 and 100. Limited discussion of the analyses leading to the conclusion that these five toxics can be quantified is given below.

Table VI.C.1-1
Toxics which can be correlated with biodiesel concentration

Acetaldehyde	% change in emissions = -0.001606 × (% biodiesel) × 100%
Ethylbenzene	% change in emissions = -0.006970 × (% biodiesel) × 100%
Formaldehyde	% change in emissions = -0.001696 × (% biodiesel) × 100%
Naphthalene	% change in emissions = -0.002847 × (% biodiesel) × 100%
Xylene	% change in emissions = -0.004078 × (% biodiesel) × 100%

Acetaldehyde: All analyses suggest a statistically significant reduction in acetaldehyde emissions with increasing biodiesel concentration. The effect is considerably smaller in magnitude than the total HC effect, which would suggest that the mass ratio of acetaldehyde to HC would increase with biodiesel concentration. This expectation is confirmed with the results in Table VI.B-2.

Ethylbenzene: All analyses suggest a statistically significant reduction in ethylbenzene emissions with increasing biodiesel concentration. The effect is slightly larger in magnitude than the total HC effect, which would suggest that the mass ratio of ethylbenzene to HC would decrease with biodiesel concentration. This expectation is confirmed with the results in Table VI.B-2.

Formaldehyde: All analyses suggest a statistically significant reduction in formaldehyde emissions with increasing biodiesel concentration. The directional result of the binomial analysis is consistent with the direction of the highly significant correlation in Table VI.B-3 even though the binomial analysis was not statistically significant.

Naphthalene: All analyses suggest a statistically significant reduction in naphthalene emissions with increasing biodiesel concentration. The directional result of the binomial analysis is consistent with the direction of the highly significant correlation in Table VI.B-3 even though the binomial analysis was not statistically significant.

Xylene: All analyses suggest a statistically significant reduction in naphthalene emissions with increasing biodiesel concentration. The effect is smaller in magnitude than the total HC effect, which would suggest that the mass ratio of ethylbenzene to HC would increase with biodiesel concentration. This expectation is not confirmed with the results in Table VI.B-2, which instead shows that the mass ratio decreases with increasing biodiesel concentration. This contrary result suggests that xylene emissions might actually exhibit reductions that are larger than the total HC reductions. However, the binomial analysis is strongly significant, indicating that xylene emissions would be expected to decrease with increasing biodiesel concentration, regardless of whether that decrease was larger or smaller in magnitude than the total HC effect. As a result, it would appear that the environmentally conservative (lowest benefit) estimate of xylene emission impacts of biodiesel use would be the correlation from Table VI.B-3.

2. Tier 2 toxics

Tier 2 toxics include those toxic compounds for which the analyses described in Section VI.B above may not be entirely consistent with one another or are not statistically significant. The effect of biodiesel on toxics cannot be quantified with confidence, but it can be estimated qualitatively in terms of directional effects. The toxics which we have determined fall into this tier include acrolein, n-hexane, and styrene. Table VI.C.2-1 gives our qualitative conclusions for these three toxics. A discussion of the analyses for these tier 2 toxics is given below.

Table VI.C.2-1
Qualitative toxic effects of biodiesel use

Acrolein	Likely small reduction
n-Hexane	Likely small reduction
Styrene	Possible increase

Acrolein: The correlation between % change in acrolein emissions and % biodiesel in Table VI.B-3 was not statistically significant, suggesting that acrolein emissions are not affected by the addition of biodiesel to conventional diesel fuel. The calculated probability in the binomial analysis would support this conclusion. However, all means for determining the direction of the effect consistently give the same answer: biodiesel reduces acrolein emissions. This includes the analyses summarized in Tables VI.B-3, VI.B-5, and VI.B-6. Table VI.B-4 shows that the combination of the acrolein/HC correlation from Table VI.B-2 with the basic HC model from Section IV.A.1 also indicates that acrolein emissions decrease with increasing biodiesel concentration. Based on these analyses we conclude that the reduction in acrolein emissions with increasing biodiesel concentration is real, but that the effect was simply too small to be captured in a statistically significant way in the correlation from Table VI.B-3.

n-Hexane: The correlation between % change in n-hexane emissions and % biodiesel in Table VI.B-3 was not statistically significant, suggesting that n-hexane emissions are not affected by the addition of biodiesel to conventional

> diesel fuel. Although the mean % change effects at 20% biodiesel and 100% biodiesel were both negative (Table VI.B-6), the value at 20% biodiesel was substantially larger in magnitude than the value at 100% biodiesel. This result may have contributed to the not significant result in Table VI.B-3. Directionally, all analyses indicate that n-hexane emissions should decrease with increasing biodiesel concentration. Table VI.B-4 shows that the combination of the n-hexane/HC correlation from Table VI.B-2 with the basic HC model from Section IV.A.1 also indicates that n-hexane emissions decrease with increasing biodiesel concentration. Based on these analyses we conclude that the reduction in n-hexane emissions with increasing biodiesel concentration is real, but that the effect was simply too small to be captured in a statistically significant way in the correlation from Table VI.B-3.
>
> *Styrene*: The correlation between % change in styrene emissions and % biodiesel in Table VI.B-3 was statistically significant, suggesting that styrene emissions are increased by the addition of biodiesel to conventional diesel fuel. However, the number of observations was extremely small, and the binomial analysis indicated no clear direction. We therefore conclude that the correlation in Table VI.B-3 can at best be used to suggest the likely direction of the effect.

3. Tier 3 toxics

Tier 3 toxics include those toxic compounds for which the analyses described in Section VI.B above are significantly in conflict with one another. The effect of biodiesel on toxics cannot be determined quantitatively or qualitatively. The toxics which we have determined fall into this tier include benzene, 1,3-butadiene, and toluene. A discussion of the analyses for these tier 3 toxics is given below.

> *Benzene*: The correlation between % change in benzene emissions and % biodiesel in Table VI.B-3 was not statistically significant, suggesting that benzene emission are not affected by the addition of biodiesel to conventional diesel fuel. Ignoring statistical significance, this correlation suggests a small increase in benzene emissions with increasing biodiesel concentration, consistent with the (highly non-significant) binomial analysis. Table VI.B-4 shows that the combination of the benzene/HC correlation from Table VI.B-2 with the basic HC model from Section IV.A.1 indicates that benzene emissions decrease with increasing biodiesel concentration, a result that is inconsistent with the trends suggested in Tables VI.B-3 and VI.B-5. In addition, the mean effect at 20% biodiesel is directionally opposite to the mean effect at 100% biodiesel (Table VI.B-6). As a result, we cannot conclude with confidence that benzene emissions are unaffected by biodiesel addition, nor can we discern a clear direction for the effect.

1,3-Butadiene: The correlation between % change in 1,3-butadiene emissions and % biodiesel in Table VI.B-3 was not statistically significant, suggesting that 1,3-butadiene emissions are not affected by the addition of biodiesel to conventional diesel fuel. Ignoring statistical significance, this correlation suggests a small decrease in 1,3-butadiene emissions with increasing biodiesel concentration, which is inconsistent with the (highly non-significant) binomial analysis. The correlation between the mass ratio of 1,3-butadiene to HC and % biodiesel was not significant, suggesting that the mass ratio should remain constant and therefore that the correlation between % change in 1,3-butadiene emissions and % biodiesel should not only have been statistically significant, but of a magnitude close to that for the basic HC model from Section IV.A. This was, however, not the case. With such conflicting results, we cannot conclude with confidence that there is no effect of biodiesel on emissions of 1,3-butadiene, nor can we infer a direction for any potential impact.

Toluene: The correlation between % change in toluene emissions and % biodiesel in Table VI.B-3 was not statistically significant, suggesting that toluene emissions are not affected by the addition of biodiesel to conventional diesel fuel. The binomial analysis would support this conclusion. However, the mean % change values at 20% biodiesel and 100% biodiesel are both positive (Table VI.B-6), suggesting that toluene emissions should increase with increasing biodiesel concentration. The correlation between the mass ratio of toluene to HC and % biodiesel was not significant, suggesting that the mass ratio should remain constant and therefore that the correlation between % change in toluene emissions and % biodiesel should not only have been statistically significant, but also negative. Table VI.B-4 shows that the combination of the toluene/HC correlation from Table VI.B-2 with the basic HC model from Section IV.A.1 indicates that toluene emissions decrease with increasing biodiesel concentration, a result that is inconsistent with the trends suggested in Table VI.B-3.

Section VII: What Additional Issues Should Be Addressed?

After reviewing the available data on emission effects of biodiesel and making efforts to quantify those effects in a robust manner, we identified those areas that could benefit from more focused attention. This Section lists those areas related to exhaust emission effects where additional work would be of value. As this list is not meant to be exhaustive, reviewers of this draft technical report are invited to suggest other areas for EPA or other stakeholders to investigate.

A. Data gaps

1. Newer highway engines

Our database contains no data on heavy-duty highway engines meeting the 2004 standards. Not only will these engines be a significant portion of the in-use fleet in the near future, but there is evidence that they may respond differently than older engines to changes in fuel properties due to the expected predominance of exhaust gas recirculation and injection rate-shaping in these newer engines. Although we attempted to evaluate how these engines might respond to biodiesel in Section IV.B.5, we cannot know for certain how these engines will respond to biodiesel without actual emissions data.

There is also some indication that injection rate-shaping has been used more commonly in engines produced in recent years. It is not known whether rate-shaping would cause engines to respond differently to biodiesel, but a combustion simulation model developed by Southwest Research Institute called ALAMO_ENGINE reportedly shows no effect of cetane on NOx emissions for engines using rate-shaping. Unfortunately, the few observations in our database representing 1998+ model years (standards group C) were collected on unique steady-state cycles, and so were not included in our analysis of heavy-duty highway engines. A preliminary review of the data on these engines does not indicate that they are likely to exhibit different responses to biodiesel than older model years. However, given the limited number of observations and the fact that the data was collected on unique steady-state test cycles, a firm conclusion cannot be reached until more data is collected.

2. Nonroad engines

Most large nonroad diesel engines use technologies similar to those found in heavy-duty highway diesel engines. On this basis we might expect biodiesel to produce similar emission effects for heavy-duty nonroad and heavy-duty highway engines. Unfortunately, as discussed in Section V.B, there were only 8 biodiesel observations collected on nonroad engines in our database. Despite the various analytical approaches we took, we could not draw a clear conclusion regarding the degree to which highway and nonroad engines respond similarly to biodiesel. In fact, there was some evidence that they respond differently. Therefore, we believe that additional data on nonroad engines should be collected before conclusions can be drawn about their emission responses to biodiesel. This would apply both to large and smaller (< 50 hp) engines.

We believe that nonroad engines very often operate in transient modes, and so may not be ideally represented by the current steady-state federal test procedure used for certification. As a result, the Agency has been in the process of developing a transient nonroad cycle in recent years. If additional biodiesel data is collected on nonroad engines, we believe that emission measurements on a transient cycle would be preferable, particularly for PM and CO.

3. Biodiesel properties

In order to more accurately determine the emission impacts of biodiesel, it would be useful to correlate actual fuel property measurements of biodiesel with emissions rather than simply using broad source categorizations as described in Section III.C.2.c. Unfortunately, as described in Section II.E.1, few of the biodiesels in our database included a full complement of compositional and physical property measurements. And yet, some properties, such as cetane number, vary considerably from one batch of biodiesel to another, while others such as specific gravity, remain relatively constant (see Figure II.E.1-1). We therefore recommend that further testing of biodiesel include more detailed property measurements of biodiesel. These properties should at minimum include cetane number, distillation, specific gravity, sulfur, aromatics, oxygen, and cloud point. Measurements of other properties would also be useful, such as lubricity characteristics or specific molecular compounds.

B. Mitigating NOx increases

One potential drawback to the use of biodiesel is the increase in NOx emissions. Although the increase is small in comparison to the reductions in other regulated pollutants, such NOx increases may be problematic for ozone nonattainment or maintenance areas. Thus additional research on ways to mitigate the NOx increase would be valuable. Various strategies might include using a lower-emitting base fuel for blending, adding a cetane improver additive to the biodiesel blend, or determining what source or properties of biodiesel can be modified to lower NOx emissions. Some of these fuel strategies have been investigated in a recent SAE paper by Robert McCormick et al[9]. Other strategies might be hardware-related, such as changing injection timing or adding a lean NOx catalyst, and additional studies of these strategies in concert with biodiesel would be valuable.

C. Base fuel effects

As discussed in Section IV.B.4, the base fuel to which biodiesel is added does appear to have an impact on the degree to which biodiesel affects exhaust emissions of regulated pollutants. There is no obvious reason for this outcome, and thus it deserves closer attention. If both the base fuels and the biodiesels in our database had a full complement of measured fuel properties, analytical means of explaining this result might be possible. Thus it would be helpful to have additional data in which the same engine is run on one batch of biodiesel blended into several different base fuels, with full speciation and property measurements for all fuels. However, a true understanding of this phenomenon might require a more in-depth study of the such things as spray atomization, diffusion burning, etc.

D. Methyl versus ethyl esters

The studies that comprised our database usually did not specify whether the biodiesel was a methyl ester or an ethyl ester. It is not clear whether this distinction is an important one, since we were unable to investigate it. It would be valuable to have information on how different transesterification processes on the same batch of plant or animal oil may lead to biodiesels exhibiting different emissions impacts.

E. Minimum data criteria

There are several aspects of our calculation of the minimum data criteria in Section III.C.1 that may warrant further investigation, since these criteria necessarily preclude the investigation of some adjustment terms in the correlations. For instance, in gasoline emissions test programs conducted by the Auto/Oil Air Quality Improvement Research Program, a determination was made that the number of vehicles was more important than the number of repeat tests on a given vehicle for accurately estimating emissions effects for the fleet. Thus it might be reasonable for this biodiesel analysis to formulate minimum data criteria that reflect engine-to-engine variability.

It might also be reasonable to use an alternative approach to calculating values for E (percent) in Table III.C.1-1 that are absolute instead of relative. In this alternative approach, equation (9) would change to the following:

$$E(\text{g/bhp-hr}) = E(\text{percent}) \times \text{Mean emissions (g/bhp-hr)}$$

where

E(g/bhp-hr) = One-half the width of the confidence interval, used in equation (8)
E(percent) = One-half the width of the confidence interval as a fixed fraction of emissions

Mean emissions = Average g/bhp-hr for conventional diesel fuel

This alternative approach would avoid the problem that the minimum data criteria for NOx were calculated to be significantly more stringent than for the other pollutants, due to the smaller relative effect of biodiesel on NOx emissions.

Appendices

Appendix A - Data Sources

Studies that were included in database

Bouché, T., M. Hinz, R. Pitterman, "Optimising Tractor CI Engines for Biodiesel Operation," SAE paper no. 2000-01-1969

Callahan, T.J., C.A. Sharp, "Evaluation of Methyl Soyate/Diesel Fuel Blends as a Fuel for Diesel Engines," Southwest Research Institute Final Report to the American Biofuels Association, December 1993

Clark, N.N., C.M. Atkinson, G.J. Thompson, R.D. Nine, "Transient Emissions Comparisons of Alternative Compression Ignition Fuels," SAE paper no. 1999-01-1117

Durbin, T.D., J.R. Collins, J.M. Norbeck, M.R. Smith, "Evalution of the Effects of Alternative Diesel Fuel Formulations on Exhaust Emissions Rates and Reactivity," Final Report from the Center for Environmental Research and Technology, University of California, April 1999

Fosseen, D., "DDC 6V-71N Emission Testing on Diesel and Biodiesel Blend," Fosseen Manufactuering and Development, Ltd., report no. NSDB4F15, July 14, 1994

Fosseen, D., "DDC 6V-92TA MU1 Coach Upgrade Emission Optimization on 20%/80% Soy/Diesel Blend," Fosseen Manufactuering & Development, Ltd., report no. 260-2 and 24 1-1, September 30, 1994

Goetz, W., "Evaluation of Methyl Soyate/Diesel Blend in a DDC 6V-92TA Engine: Optimization of NOx Emissions," Ortech International, Report No. 93-E14-36, July 20, 1993

Goetz, W., "Evaluation of a Tallow/Diesel Blend in a DDC 6V-92TA Engine," Ortech International report no. 93-E14-37, July 21, 1993

Graboski, M.S., J.D. Ross, R.L. McCormick, "Transient Emissions from No. 2 Diesel and Biodiesel Blends in a DDC Series 60 Engine," SAE paper no. 961166

Graboski, M.S., R.L. McCormick, T.L. Alleman, A.M. Herring, "The Effect of Biodiesel Composition on Engine Emissions from a DDC Series 60 Diesel Engine," Colorado School of Mines, Final Report to National Renewable Energy Laboratory, June 8, 2000

Hansen, K.F., M.G. Jensen, "Chemical and Biological Characteristics of Exhaust Emissions from a DI Diesel Engine Fuelled with Rapeseed Oil Methyl Ester (RME)," SAE paper no. 971689

Howes, P., G. Rideout, "Evaluation of Biodiesel in an Urban Transit Bus Powered by a 1988 DDECII6V92 TA Engine," National Biodiesel Board, MSED Report # 96-26743-1, June 1995

Howes, P., G. Rideout, "Evaluation of Biodiesel in an Urban Transit Bus Powered by a 1981 DDCC8V71 Engine," National Biodiesel Board, MSED Report # 95-26743-2

Liotta, F.J., D.M. Montalvo, "The Effect of Oxygenated Fuels on Emissions from a Modern Heavy-Duty Diesel Engine," SAE paper no. 932734

Manicom, B., C. Green, W. Goetz, "Methyl Soyate Evaluation of Various Diesel Blends in a DDC 6V-92 TA Engine," Ortech International, April 21, 1993

Marshall, W., L.G. Schumacher, S. Howell, "Engine Exhaust Emissions Evaluation of a Cummins L10E When Fueled With a Biodiesel Blend," University of Missouri.

McCormick, R.L., J.D. Ross, M.S. Graboski, "Effect of Several Oxygenates on Regulated Emissions from Heavy-Duty Diesel Engines," *Environmental Science and Technology,* 1997*, 31,* 1144-1150

McCormick, R.L., J.R. Alvarez, M.S. Graboski, "NOx Solutions for Biodiesel," Colorado School of Mines, August 31, 2001

McDonald, J.F., D.L. Purcell, B.T. McClure, D.B. Kittelson, "Emissions Characteristics of Soy Methyl Ester Fuels in an IDI Compression Ignition Engine," SAE paper no. 950400

Peterson, C., D. Reece, J. Thompson, S. Beck, C. Chase, "Development of Biodiesel for Use in High Speed Diesel Engines," University of Idaho, presentation at Sixth National Bioenergy Conference, Oct. 2-6, 1994.

Peterson, C.L., D.L. Reece, "Emissions Tests with an On-Road Vehicle Fueled with Methyl and Ethyl Esters of Rapeseed Oil," presented at the 1994 ASAE International Winter Meeting, ASAE paper no. 946532.

Peterson, C.L., D.L. Reece, "Emissions Testing with Blends of Esters of Rapeseed Oil Fuel With and Without a Catalystic Converter," SAE paper no. 961114

Peterson, C.L., "Truck-In-The-Park Biodiesel Demonstration with Yellowstone National Park," University of Idaho, August 1999.

Rantanen, L., S. Mikkonen, L. Nylund, P. Kociba, M. Lappi, N. Nylund, "Effect of Fuel on the Regulated, Unregulated and Mutagenic Emissions of DI Diesel Engines," SAE paper no. 932686

Schumacher, L.G., S.C. Borgelt, W.G. Hires, D. Fosseen, W. Goetz, "Fueling Diesel Engines with Blends of Methyl Ester Soybean Oil and Diesel Fuel," University of Missouri. See http://web.missouri.edu/~pavt0689/ASAED94.htm.

Schumacher, L., S.C. Borgelt, W.G. Hires, W. Wetherell, A. Nevils, "100,000 Miles of Fueling 5.9L Cummins Engines with 100% biodiesel," SAE paper no. 962233

Sharp, C.A., "Transient Emissions Testing of Biodiesel in a DDC 6V-92TA DDEC Engine," Southwest Research Institute, Final Report to the National Biodiesel Board, October 1994

Sharp, C.A., "Transient Emissions Testing of Biodiesel and Other Additives in a DDC Series 60 Engine," Southwest Research Institute, Final Report to the National Biodiesel Board, December 1994

Sharp, C.A., "Emissions and Lubricity Evaluation of Rapeseed Derived Biodiesel Fuels," Southwest Research Institute, Final Report for Montana Department of Environmental Quality, November 1996

Sharp, C.A., "Characterization of Biodiesel Exhaust Emissions for EPA 211(b)," Southwest Research Institute report no. 08-1039A, January 1998

Sharp, C.A., S.A. Howell, J. Jobe, "The Effect of Biodiesel Fuels on transient Emissions from Modern Diesel Engines, Part I Regulated Emissions and Performance," SAE paper no. 2000-01-1967

Sirman, M.B., E.C. Owens, K.A. Whitney, "Emissions Comparison of Alternative Fuels in an Advanced Automotive Diesel Engine," Southwest Research Institute, AD A353968/PAA, November 1998

Smith, J.A., D.L. Endicott, R.R. Graze, "Biodiesel Engine Performance and Emissions Testing," Caterpillar Technical Center, May 1998

Spataru, A., C. Romig, "Emissions and Engine Performance from Blends of Soya and Canola Methyl Ester with ARB #2 Diesel in a DDC 6V92TA MUI Engine," SAE paper no. 952388

Starr, M.E., "Influence on Transient Emissions at Various Injection Timings, Using Cetane Improvers, Bio-Diesel, and Low Aromatic Fuels," SAE paper no. 972904

Stotler, R.W., D.M. Human, "Transient Emission Evaluation of Biodiesel Fuel Blend in a 1987 Cummins L10 and DDC 6V-92-TA," Engineering Test Services, Report No. ETS-95-128, Nov. 30, 1995

Ullman, T.L., C.T. Hare, T.M. Baines, "Heavy-Duty Diesel Emissions as a Function of Alternate Fuels," SAE paper no. 830377

"Emissions from Biodiesel Blends and Neat Biodiesel from a 1991 Model Series 60 Engine Operating at High Altitude," Colorado Institute for Fuels and High Altitude Engine Research, Colorado School of Mines, September 1994.

"Effects of Methyl Esters of Tallow and Grease on Exhaust Emissions and Performance of a Cummins L10 Engine," National Institute for Petroleum and Energy Research, September 16, 1993

Studies that were not included in the database because the data was collected on an engine that was single cylinder or otherwise experimental

Murayama, T., Y. Oh, A. Kido, T. Chikahisa, N. Miyamoto, K. Itow, "Effects of Super Heating of Heavy Fuels on Combustion and Performance in DI Diesel Engines," SAE paper no. 860306

Murayama, T., Y. Oh, N. Miyamoto, T. Chikahisa, N. Takagi, K. Itow, "Low Carbon Flower Buildup, Low Smoke, and Efficient Diesel Operation with Vegetable Oils by Conversion to Mono-Esters and Blending with Diesel Oil or Alcohols," SAE paper no. 841161

Shaheed, A., E. Swain, "Performance and Exhaust Emission Evaluation of a Small Diesel Engine Fuelled with Coconut Oil Methyl Esters," SAE paper no. 981156

Suda, K.J., "Vegetable Oil or Diesel Fuel - A Flexible Option," SAE paper no. 840004

Ziejewski, M., H.J. Goettler, "Comparative Analysis of the Exhaust Emissions for Vegetable Oil Based Alternative Fuels," SAE paper no. 920195

Studies that were not included in the database because the emissions data was not readily available

Alfuso. S., M. Auriemma, G. Police, M.V. Prati, "The Effect of Methyl-Ester of Rapeseed Oil on Combustion and Emissions of DI Diesel Engines," SAE paper no. 932801

Hemmerlain, N., V. Korte, H. Richter, G. Schröder, "Performance, Exhaust Emissions and Durability of Modern Diesel Engines Running on Rapeseed Oil," SAE paper no. 910848

Humke, A.L., N.J. Barsic, "Performance and Emissions Characteristics of a Naturally Aspirated Diesel Engine with Vegetable Oil Fuels - (Part 2)," SAE paper no. 810955

Last, R.J., M. Krüger, M. Dürnholz, "Emissions and Performance Characteristics of a 4-Stroke, Direct Injected Diesel Engine Fueled with Blend of Biodiesel and Low Sulfur Diesel Fuel," SAE paper no. 950054

Ma, W., L.I. Leviticus, F.G. Ullman, "On-Line Measurement of Formaldehyde in Tailpipe Emissions by Tunable Diode Laser Spectroscopy," SAE paper no. 941702

Martin, B., P. Aakko, D. Beckman, N. D. Giacomo, F. Giavazzi, "Influence of Future Fuel Formulations on Diesel Engine Emissions - A Joint European Study," SAE paper no. 972966

McDonald, J.F., "Evaluation of a Yellow Grease Methyl Ester and Petroleum Diesel Fuel Blend," University of Minnesota Final Report to the Agricultural Utilization Research Institute, August 11, 1997

Montagne, X., "Introduction of Rapeseed Methyl Ester in Diesel Fuel - The French National Program," SAE paper no. 962065

Reece, D.L., C.L. Peterson, "Biodiesel Testing in Two On-Road Pickups," SAE paper no. 952757

Staat, F., P. Gateau, "The Effects of Rapeseed Oil Methyl Ester on Diesel Engine Performance, Exhaust Emissions and Long-Term Behavior - A Summary of Three Years of Experimentation," SAE paper no. 950053

Wang, W.G., D.W. Lyons, N.N. Clark, M. Gautam, "Emissions from Nine Heavy Trucks Fueled by Diesel and Biodiesel Blend without Engine Modification," *Environmental Science & Technology*, **2000**, *34*, 933-939

Studies that were not included in the database because the test cycle only included a small number of nonstandard modes or was otherwise unrepresentive of the FTP

Adelman, A.J., "Emission Evaluation Test Report for a Comparative Analysis of Emissions Among Petroleum Diesel and Biodiesel Blends fired in a Large Diesel Engine," AirNova, Inc., January 1998

Akasaka, Y., T. Suzuki, Y. Sakurai, "Exhaust Emissions of a DI Diesel Engine Fueled with Blends of Biodiesel and Low Sulfur Diesel Fuel," SAE paper no. 972998

Bagley, S.T., L.D. Gratz, J.H. Johnson, J.F. McDonald, "Effects of an Oxidation Catalytic Converter and a Biodiesel Fuel on the Chemical, Biological, and Particulate Size Characteristics of Emissions from an IDI Diesel Engine," Michigan Technological University, December 1995

Chang, D.Y.Z., J.H. Van Gerpen, "Fuel Properties and Engine Performance for Biodiesel Prepared from Modified Feedstocks," SAE paper no. 971684

Chang, D.Y., J.H. Van Gerpen, "Determination of Particulate and Unburned Hydrocarbon Emissions from Diesel Engines Fueled with Biodiesel," SAE paper no. 982527

Choi, C.Y., G.R. Bower, R.D. Reitz, "Effects of Biodiesel Blended Fuels and Multiple Injections on D.I. Diesel Engines," SAE paper no. 970218

Czerwinski, J., "Performance of HD-DI-Diesel Engine with Addition of Ethanol and Rapeseed Oil," SAE paper no. 940545

Desantes, J.M., J. Arrègle, S. Ruiz, A. Delage, "Characterization of the Injection-Combustion Process in a D.I. Diesel Engine Running with Rape Oil Methyl Ester," SAE paper no. 1999-01-1497

Fort, E.F., P.N. Blumberg, H.E. Staph, J.J. Staudt, "Evaluation of Cottonseed Oil as Diesel Fuel," SAE paper no. 820317

Jacobus, M.J., S.M. Geyer, S.S. Lestz, W.D. Taylor, T.H. Risby, "Single-Cylinder Diesel Engine Study of Four Vegetable Oils," SAE paper no. 831743

Masjuki, H., M.Z. Abdulmuin, H.S. Sii, "Investigations on Pre-Heated Palm Oil Methyl Esters in the Diesel Engine," IMechE 1996, Part A: Journal of Power and Energy

Needham, J.R., D.M. Doyle, "The Combustion and Ignition Quality of Alternative Fuels in Light Duty Diesels," SAE paper no. 852101

Schmidt, K., J. Van Gerpen, "The Effect of Biodiesel Fuel Composition on Diesel Combustion and Emissions," SAE paper no. 961086

Scholl, K.W., S.C. Sorenson, "Combustion of Soybean Oil Methyl Ester in a Direct Injection Diesel Engine," SAE paper no. 930934

Schramm, J., I. Foldager, N. Olsen, L. Gratz, "Emissions from a Diesel VehicleOperated on Alternative Fuels in Copenhagen," SAE paper no. 1999-01-3603

Schröder, O., J. Krahl, A. Munack, J. Krahl, J. Bünger, "Environmental and Health Effects Caused By the Use of Biodiesel," SAE paper no. 1999-01-3561

Senatore, A., M. Cardone, V. Rocco, M.V. Prati, "A Comparative Analysis of Combustion Process in D.I. Diesel Engine Fueled with Biodiesel and Diesel Fuel," SAE paper no. 2000-01-0691

Uchida, M., Y. Akasaka, "A Comparison of Emissions from Clean Diesel Fuels," SAE paper no. 1999-01-1121

Zhang, Y., J.H. Van Gerpen, "Combustion Analysis of Esters of Soybean Oil in a Diesel Engine," SAE paper no. 960765

Studies that were not included in the database because they did not contain any new emissions data for regulated pollutants

Culshaw, F.A., "The Potential of 'Biodiesel' from Oilseed Rape," IMechE 1993, Part A: Journal of Power and Energy

Durbin, T.D., J.R. Collins, H. Galdamez, J.M. Norbeck, M.R. Smith, R.D. Wilson, T. Younglove, "Evaluation of the Effects of Biodiesel Fuel on Emissions from Heavy-Duty Non-Road Engines," Univerity of California Final Report submitted to South Coast Air Quality Management District, May 2000

Peterson, C.L., D.L. Auld, "Technical Overview of Vegetable Oil as a Transportation Fuel," University of Idaho, 1991

Sharp, C.A., S.A. Howell, J. Jobe, "The Effect of Biodiesel Fuels on Transient Emissions from Modern Diesel Engines, Part II Unregulated Emissions and Chemical Characterization," SAE paper no. 2000-01-1968

Studies that were not included in the database for other reasons

Watts, W.F., M. Spears, J Johnson, "Evaluation of Biodiesel Fuel and Oxidation Catalyst in an Underground Metal Mine," University of Minnesota, September 24, 1998

> *Reason: Emission measurements were made in the field with portable emission measurement devices*

Yoshimoto, Y., M. Onodera, H. Tamaki, "Reduction of NOx, Smoke, and BSFC in a Diesel Engine Fueled by Biodiesel Emulsion with Used Frying Oil," SAE paper no. 1999-01-3598

> *Reason: Biodiesel blend included water in an emulsion*

Appendix B - Field descriptions for database

Table	Field	Definition
EQUIP_AD	equip_id	Unique mobile source identifier. Serial number is ideal. Otherwise some variant of the study_id.
	study_id	Identification number assigned to the analysis/paper/report of interest.
	Class	Light-duty or heavy-duty
	equip_type	Vehicle or engine
	Company	Vehicle or engine manufacturer.
	highway	"Yes" if mobile source is intended for highway use. "No" for non-road mobile sources.
	model_name	model name
	model_yr	If a prototype, enter representative model year.
	make	Vehicle make e.g. Buick, as distinct from vehicle manufacturer, GM.
	disp_liter	Nominal engine displacement, expressed in liters.
	fi_type	type of fuel injection PFI (port fuel injection) TBI (throttle body injection) INDIR (Indirect injection) DIRECT (direct fuel injection e.g. as for most diesel engines.) Values defined by translation table for this field.
	aspirated	indicates how engine is aspirated. CHARGED if turbocharged or supercharged. NATURAL if not. Values defined by translation table for this field.
	cylinder	Number of cylinders or rotors.
	cat_type	What type catalyst, if any, is present on the mobile source. Values are: 3WAY Three-way catalyst OX3W Oxidation plus three-way catalyst OXID Oxidation Catalyst NONE No catalyst NULL Unknown Values defined by translation table for this field.
	egr_type	Type of exhaust gas recirculation (EGR). Values defined by translation table. Values defined by translation table for this field.
	engseries	Engine series or product line name.
	cooling	Type of after_cooling. (Legal values defined by translation table.) Values defined by translation table for this field.

	fi_meth	Method of fuel injection. (Legal values defined by translation table.)
	fi_press	Fuel injection pressure, expressed in kPa.
	parttrap	Is particulate trap used? "YES", "NO", or "NUL".
	eng_cycle	Engine cycle, 2 =. 2-stroke, 4 = 4-stroke. 0 = Unknown. Values defined by translation table for this field.
	ratedpower	Rated horsepower of engine.
	ratedspeed	Rated rpm of engine
	idle_rpm	Idle rpm as declared by the oem.
	proc_odom	Approximate odometer reading in miles at time of vehicle recruitment.
	hour_meter	Hours of operation (usually available only for off-road mobile sources). Null value is 0.
	gvwr	Gross vehicle weight rating in pounds. The value specified by the manufacturer as the loaded weight of a single vehicle.
	pk_torque	Peak torque of the engine expressed in ft-lb.
	pk_t_speed	Peak torque speed expressed in rpm.
	cyl_valves	The number of valves per cylinder.
	stroke	Piston stroke expressed in inches. (not ready to be stored in msod database at this time)
	bore	The diameter of the cylinder expressed in inches.
	inj_ctrl	Code of the Injection control type. Values defined by translation table for this field.
	inj_pcat	Code of the injection equipment/pressure category. Values defined by translation table for this field.
ETEST_AD	test_id	Identification number assigned to the engine test.
	study_id	Identification number assigned to the analysis/paper/report of interest.
	fbatch_id	Fuel batch identification.
	equip_id	Unique mobile source identifier. Serial number is ideal. Otherwise some variant of the study_id.
	test_proc	Identifies the specific test procedure/cycle used. Values defined by translation table for this field. "Steady-State" if a unique steady-state cycle, "Transient" for unique transient cycle.
	No_modes	For steady-state tests, the number of modes in the test

	ch4	Methane emissions. Expressed in grams per bhp-hr for engines, and g/mi for vehicles.
	thc	Total HC emissions. Expressed in grams per bhp-hr for engines, and g/mi for vehicles.
	co	CO emissions. Expressed in grams per bhp-hr for engines, and g/mi for vehicles.
	nox	NOx emissions. Expressed in grams per bhp-hr for engines, and g/mi for vehicles.
	pm	Total particulate emissions. Expressed in grams per bhp-hr for engines, and g/mi for vehicles.
	total_work	Total work performed in test. Expressed in bhp-hrs.
	bsfc_meas	Measured brake-specific fuel consumption. Expressed in grams per bhp-hr for engines, and g/mi for vehicles.
FBAT_AD	fbatch_id	Unique fuel batch identification.
	fbatch_base	Unique identification linking all batches that are intended to be compared to one another.
	study_id	Identification number assigned to the analysis/paper/report of interest.
	cetane_num	Total cetane number of complete fuel.
	cetane_idx	Cetane index of complete fuel.
	cetane_imp	Amount of cetane improver added, expressed as percentage by volume
	cetane_typ	Type of cetane improver used, e.g. "N" for nitrate type or "P" for peroxide type. Exact set of legal values defined and described by translation table for this field.
	cetane_dif	This is the difference in cetane number between the described fuel (with additive) and a baseline fuel without additive.
	cetane_nat	Natural cetane number of fuel.
	sulfur	Sulfer content, expressed in parts per million.
	nitrogen	Nitrogen content, expressed in parts per million.
	tarom	Total aromatics content of fuel, expressed as a percentage by volume. This is a measured value, as opposed as being calculated as the sum of the monoaromatics and polyaromatics fields.
	marom	Monoaromatics content of fuel, expressed as a percentage by weight. This is a measured value, as opposed as being

		calculated as the difference of the total aromatics and polyaromatics fields.
	parom	Polyaromatics content of fuel, expressed as a percentage by weight. This is a measured value, as opposed as being calculated as the difference of the total aromatics and monoaromatics fields.
	IBP	Initial boiling point expressed in degrees F.
	T10	10% distillation boiling point, expressed in degrees Fahrenheit.
	T50	50% distillation boiling point, expressed in degrees Fahrenheit.
	T90	90% distillation boiling point, expressed in degrees Fahrenheit.
	T95	95% distillation boiling point, expressed in degrees Fahrenheit.
	EP	End point of distillation curve, expressed in degrees Fahrenheit.
	spec_grav	Specific gravity.
	viscosity	Viscosity, expressed in centistokes @40 degrees F.
	hcratio	Molecular ratio of hydrogen to carbon.
	oxygen	Amount of oxygen in the fuel, expressed as a percentage by weight.
	bio_source	Source of biodiesel. "NONE" if no biodiesel was added. Values defined by translation table for this field.
	bio_type	Type of biodiesel. "Oil" or "Ester"
	bio_cat	Category of biodiesel. "Plant" or "animal"
	per_bio	Volume percent biodiesel in the blend
	heat	Net heating value of the fuel, expressed in btu/pound.
	emiss_cat	Emission category applicable to base fuel. "Clean" or "average"
	Bio_source_grp	Biodiesel source group. "Fat", "soy", or "rape". "Base" applies to base fuels.

Appendix C - Assignments for biodiesel source groups

All base fuels were assigned to the "average" emissions category, except for those in the following table which were assigned to the "clean" category.

Study	Fuel	EMISS_CAT assignment
SAE 932686 (Rantanen 1993)	SF2	Clean
SAE 971689 (Hansen 1997)	Ultra-light diesel	Clean
Durbin 1999	RFD	Clean
Manicom 1993	Esso diesel #1	Clean
McCormick 2001	10 vol% aromatics	Clean
McCormick 2001	Fischer-Tropsch	Clean
McCormick 2001	No. 1 diesel	Clean
Sirman 1998	Low sulfur	Clean

Appendix D - Studies used in toxics analysis

Durbin, T., J. Collins, J. Norbeck, and M. Smith. "Evaluation of the Effects of Alternative Diesel Fuel Formulations on Exhaust Emission Rates and Reactivity," Center for Environmental Research and Technology, University of California. April 1999

Howes, P. and G. Rideout. "Evaluation of Biodiesel in an Urban Transit Bus Powered by a 1981 DDC8V71 Engine," MSED Report #95-26743-2. 1981.

Howes, P. and G. Rideout. "Evaluation of Biodiesel in an Urban Transit Bus Powered by a 988 DDECII6V92 TA Engine," MSED Report #95-26743-1. 1988.

Sharp, C. "Transient Emissions Testing of Biodiesel and Other Additives in a DDC Series 60 Engine," Prepared for National Biodiesel Board by Southwest Research Institute. December 1994.

Sharp, C. "Emissions and Lubricity Evaluation of Rapeseed Derived Biodiesel Fuels," Prepared for Montana Department of Environmental Quality by Southwest Research Institute. November 1996.

Sharp, C. Characterization of Biodiesel Exhaust Emissions for EPA 211(b). Prepared for National Biodiesel Board by Southwest Research Institute. January 1998.

Sharp, C., S. Howell, and J. Jobe. "The Effect of Biodiesel Fuels on Transient Emissions from Modern Diesel Engines, Part II Unregulated Emissions and Chemical Characterization," 2000.

Staat, F. and P. Gateau. "The Effects of rapeseed Oil Methyl Ester on Diesel Engine Performance, Exhaust Emissions and Long-Term Behavior - A Summary of Three Years of Experimentation," No date.

Appendix E - Aromatics Conversion Equations

The conversion equation described in this appendix were originally derived and presented in Appendix C of the July 2001 Staff Discussion Document. The reader is refered to that document for details of the derivations.

Table E1 - Correlations for total aromatics

[vol% by FIA] = 0.777 × [wt% by mass spec] + 132.2 × [sp. gravity] - 105.0	$R^2 = 0.93$
[vol% by FIA] = 0.760 × [wt% by HPLC] + 178.0 × [sp. gravity] - 144.4	$R^2 = 0.96$

Table E2 - Correlations for mono aromatics

[wt% by SFC] = 0.882 × [wt% by mass spec] + 2.37	$R^2 = 0.91$
[wt% by SFC] = 0.885 × [wt% by HPLC] + 0.88	$R^2 = 0.99$

Table E3 - Correlations for poly aromatics

[wt% by SFC] = 1.22 × [wt% by mass spec] + 0.33	$R^2 = 0.95$
[wt% by SFC] = 1.27 × [wt% by HPLC] + 0.69	$R^2 = 0.97$

References

1. "Strategies and Issues in Correlating Diesel Fuel Properties with Emissions," Staff Discussion Document, EPA document number EPA420-P-01-001, July 2001

2. R. Lee, J. Pedley, C. Hobbs, "Fuel Quality Impact on Heavy Duty Diesel Emissions: -A Literature Review", SAE 982649.

3. "Strategies and Issues in Correlating Diesel Fuel Properties with Emissions," Staff Discussion Document, EPA document number EPA420-P-01-001, July 2001

4. *Statistics Today: A Comprehensive Introduction*, Donald R. Byrkit, 1987, The Benjamin/Cummings Publishing Company, Inc. Page 321

5. "Strategies and Issues in Correlating Diesel Fuel Properties with Emissions," Staff Discussion Document, EPA document number EPA420-P-01-001, July 2001

6. Rickeard, D.J., N.D. Thompson, "A Review of the Potential for Bio-Fuels as Transportation Fuels," SAE paper number 932778

7. Sheehan, J., V. Camobreco, J. Duffield, M. Graboski, H. Shapouri, "Life Cycle Inventory of Biodiesel and Petroleum Diesel for Use in an Urban Bus," Final Report, National Renewable Energy Laboratory, May 1998, NREL/SR-580-24089

8."Control of Emissions of Hazardous Air Pollutants from Mobile Sources," 66 FR 17230, March 29, 2001

9. McCormick, R.L., J.R. Alvarez, M.S. Graboski, K.S. Tyson, K. Vertin, "Fuel additive and blending approaches to reducing NOx emissions from biodiesel," SAE paper no. 2002-01-1658.