Peer Review of *Assessing the Effect of Five Gasoline Properties on Exhaust Emissions from Light-Duty Vehicles certified to Tier-2 Standards (EPAct/V2/E-89: Phase 3)*

Part II: Data Analysis and Model Development

Prepared for

Assessment and Standards Division

Office of Transportation and Air Quality

U.S. Environmental Protection Agency

Prepared by

Systems Research and Application Corporation

652 Peter Jefferson Parkway, Suite 300

Charlottesville, VA 22911

EPA Contract Number: EP-C-11-007

February 2012

Peer Review of *Assessing the Effect of Five Gasoline Properties on Exhaust Emissions from Light-Duty Vehicles certified to Tier-2 Standards (EPAct/V2/E-89: Phase 3)*

Part II: Data Analysis and Model Development

Table of Contents

TO: Kent Helmer, Connie Hart, U.S. Environmental Protection Agency, Office of Transportation and Air Quality (OTAQ)

FROM: Brian Menard, SRA International

DATE: February 10, 2012

SUBJECT: Peer Review of *Assessing the Effect of Five Gasoline Properties on Exhaust Emissions from Light-Duty Vehicles certified to Tier-2 Standards (EPAct/V2/E-89: Phase 3) – Part II: Data Analysis and Model Development*

1. Background

Past fuel effects models, such as the EPA Predictive Model and EPA's Complex Model, were based on data collected using 1990s-technology cars and trucks meeting Tier 0 and Tier 1 vehicle emission standards at pollutant levels an order of magnitude higher than compliance levels for current Tier2 vehicles. With the current on-highway fleet turning out much lower-emitting vehicles than in past years, the U.S. Congress, the EPA and a variety of stakeholders are interested in generating an updated data set for fuel effects models to guide vehicle emissions policy. Further, Section 1506 of the Energy Policy Act (EPAct) of 2005 directed EPA to generate a more current fuel effects model representing future gasoline vehicle fleet emissions. The anti-backsliding studies requested by Section 211(q) and 211(v) of the Clean Air Act would benefit from the output of updated fuel effects models.

This report, *Assessing the Effect of Five Gasoline Properties on Exhaust Emissions from Light-Duty Vehicles certified to Tier-2 Standards (EPAct/V2/E-89: Phase 3) – Part II: Data Analysis and Model Development (EPAct Study Analysis)*, describes the analysis and modeling of data collected in Phase 3 of the EPAct/V2/E-89 light duty gasoline vehicle fuel effects study (referred to here as "the EPAct program", or "Phase 3 of the EPAct program"). This study examined the exhaust emission impacts on vehicles of changing levels of five fuel properties (ethanol, T50, T90, aromatics, and RVP (specified as DVPE)) in a matrix of 27 gasoline fuels, developed to implement an optimal design allowing estimation of 11 effects (5 linear effects, 2 quadratic effects and four 2-way interactions). The range in the levels of fuel properties tested spans the ranges expected in current market fuels and in potential mid-level ethanol blended fuels. The test vehicle sample consisted of 15 new 2008 light-duty cars and trucks, selected from among high sales makes and models, to provide a representative sample of a fleet of vehicles meeting the U.S. Federal Tier 2 vehicle emission standards. Given the relatively low level of emissions from these vehicles, a number of design and procedural steps were undertaken to minimize the impacts of measurement variability and other artifacts on data quality.

Vehicle emission test data collected in the program included concentrations and rates for typical regulated pollutants on an aggregate (bag) and continuous (second-by-second) basis, plus data for speciated emissions for subsets of vehicles and fuels. Complete data was generated for 926 tests, with 30 additional tests containing valid measurements for regulated emissions. The analysis of this data focuses on producing reduced mixed-effect models by emission bag for the pollutants of interest (nitrogen oxides, total hydrocarbon, non-methane organic gases and carbon monoxide) after screening for outliers and other data quality issues. Fuel effects models were developed starting with the set of five linear effects and the target set of six 2^{nd}-order terms for which the fuel matrix was optimized (total of 11 terms), plus a set of additional 2^{nd-}order terms comprising possible remaining interactive terms (for a total of 16 terms) and evaluating fit parameters for both sets of models. A conservative approach was taken on outlying observations, such that statistical parameters outside specified thresholds were generally not sufficient to remove data without identification of an underlying measurement or procedural problem. Relatively few outliers were removed on the basis that their values were unrealistically high (limited to four measurements, all from the particulate mass emission dataset), but subsets of data representing particular vehicles were removed after data quality review suggested some very low-level measurements from these vehicles substantially affected by measurement error. Another issue encountered during analysis was treatment of non-detects (zero values), which were addressed via use of Tobit regressions (employing the PROC LIFEREG procedure in SAS).

The primary objective of this program is to generate data to produce models with good predictive abilities for vehicles and fuels beyond those specifically tested in this program. Thus, minimizing retention of terms that are based largely on measurement or design artifact or are otherwise overfit to these vehicles and fuels is a concern.

EPA sought an expert peer review of the *EPAct Study Analysis*, including reviewers' opinion on the appropriateness of the statistical techniques described in the report and their appropriateness in the context of data accuracy/quality issues. This report documents the peer review. Section 2 of this memorandum describes the process for selecting reviewers, administering the review process, and closing the peer review. Section 3 summarizes reviewer comments according to the series of specific questions set forth in the peer review charge. The appendices to the memorandum contain the peer reviewers' resumes, completed conflict of interest and bias questionnaires for each reviewer, and the peer review charge letter.

2. Description of Review Process

In October 2011, OTAQ contacted SRA International to facilitate the peer review of the *EPAct Study Analysis*. EPA provided SRA with a list of subject matter experts from academia, consulting, and industry to serve as a starting point for identification of peer reviewer candidates. SRA selected three independent (as defined in Sections 1.2.6 and 1.2.7 of EPA's *Peer Review Handbook, Third Edition)* subject matter experts to conduct the requested reviews. SRA selected subject matter experts familiar with statistical analysis and vehicle emissions. To ensure the independence and impartiality of the peer review, SRA was solely responsible for selecting the peer review panel. Appendix A of this report contains the resumes of the three peer reviewers. A crucial element in selecting peer reviewers was to determine whether reviewers had any actual or perceived conflicts of interest or bias that might prevent them from conducting a fair and impartial review of the *EPAct Study Analysis*. SRA required each reviewer to complete and sign a conflict of interest and bias questionnaire. Appendix B of this report contains an explanation of the process and standards for judging conflict and bias along with copies of each reviewer's signed questionnaire.

SRA provided the reviewers a copy of the most recent version of the *EPAct Study Analysis* as well as the peer review charge containing specific questions EPA asked the reviewers to address. Appendix C of this report contains the memo to reviewers from SRA with the peer review charge.

SRA delivered the final review comments to EPA by the requested date. These reviews, contained in Appendix D of this report, include the reviewers' response to the specific charge questions and any additional comments they might have had.

3. Compilation of Review Comments

The *EPAct Study Analysis* was reviewed by Dr. Xuming He (University of Michigan), Dr. Christian Lindhjem (ENVIRON), and Mr. Brian West (Oak Ridge National Laboratory). Appendix A contains detailed resumes for each of the reviewers. This section is a compilation of their comments. The comments have been categorized as specific, general, and editorial. All textual edits and corrections provided by the three reviewers are indicated in bold type. The reviewers' complete comments may be found in Appendix D.

3.1 Specific Technical Comments

The reviewers provided a significant number of specific technical comments and were generally favorable in their reviews of the technical and statistical aspects of the *EPAct Study Analysis*. This section contains their comments and is divided into those that specifically address either questions or requests contained in the peer review charge and additional technical comments that reviewers chose to provide.

3.1.1 Charge Questions & Requests for Comment

In addition to encouraging reviewers to best apply their particular area(s) of expertise to review the overall study, EPA drafted six questions or requests for comment to serve as a focus for the reviewers. These were included in the peer review charge provided to the reviewers. In varying degree, the three reviewers provided direct responses to the questions or requests for comments.

1. **Was the process of imputation of NMOG/NMHC results for tests/bags with missing speciation data reasonable and statistically sound? (Section 2.2)**

He: Due to the speciation schedule described in Section 2.2 of the report, most tests in the dataset do not have alcohol and carbonyl measurements for bags 2 and 3. As NMOG and NMHC are calculated emission results that use speciation data, they could not be computed for the portions of the dataset without speciation. The study used imputation based on an alternate measure of hydrocarbon emissions to fill in the missing values. Linear location-scale-type models were used to imputation with special substitution of value zero for small NMOG. The report made a convincing case that the models used for imputation fit the data well, and should result in small errors and variability due to imputation.

It seems that the imputed values were deterministic in the study given the variables *x*NMHC in Equations 8-11. If so, statistical variability could be under-reported in the subsequent studies. One approach to recommend here is to use multiple imputations to account for the variability. Based on the descriptions in the report, I do not think that the additional variability due to imputation would be a significant factor, but I would prefer to see a more explicit discussion and examination of this issue in the study.

Lindhjem: For example, this statement on Page 36 ("*The alternate measure "NMHC as measured by FID" ($NMHC_{FID}$), was collected for the entire dataset, and it very* ***tightly*** *correlated with both NMOG and "true" NMHC. It is thus possible to estimate NMOG and NMHC results for tests without speciation by using correlations generated from those with speciation. This technique essentially estimates the offset between the response of the FID and the fully characterized emission stream, due to the incomplete measurement of oxygenates by the FID. For NMOG, this estimated value is typically between 2-20% higher than the $NMHC_{FID}$ measurement, depending on emission bag and fuel ethanol level."*) refers to an apparently unique measurement $NMHC_{FID}$ and data handling approach, and it would be useful to understand this measurement in order to understand the validity of this statement. The correlation (described in section 3.2 Imputation of Speciated Hydrocarbons (NMOG, NMHC)) of $NMHC_{FID}$ to NMOG and NMHC appears to indicate that fuel ethanol level has no effect on this correlation (same slope for all levels of ethanol, Tables 10, 11 & 13, 14) for Bag 2 and 3, but an ethanol slope term for Bag 1. Yet, the correlations of acetaldehyde, formaldehyde, and ethanol in Table 82 for Bag 2 demonstrate ethanol still increases these oxygenated species and so should increase NMOG (consisting usually of NMHC + oxygenated carbon). Perhaps there is something unique about this $NMHC_{FID}$ measurement that allows oxygenated species to be measured at some level. However, the approach to correlating NMOG and NMHC with $NMHC_{FID}$ appears inconsistent for Bags 2 and 3 compared to Bag 1.

West: Estimating NMOG from available NMHC appears reasonable and is consistent with what was done in the DOE V4 program (ORNL/TM-2011/234 and ORNL/TM-2011/461). For a given ethanol level, NMOG emissions have been shown to be a linear function of NMHC emissions.

2. **Was the decision to remove very low emitting and influential vehicles from the NOx and NMOG analyses reasonable? (Sections 5.5 and 6.1)**

He: Modifying or removing outliers and influential observations could raise questions about the validity of a study, especially when ad hoc decisions are made after the data are collected and examined. The report paid serious attention to data quality, and described how outliers and influential observations were identified.

At the end of Section 5.2, it was mentioned that Run 6281 in Bag 3 was removed even though it was not flagged as influential. The specific reason for this decision was lacking in the report.

West: These discussions were largely reasonable and convincing; however one case (run 6281) was not explained.

3. **Please comment on the use of the "design set" of 11 terms as the basis of the final models, versus allowing model to fit all 17 terms, including the adequacy of the justification of this decision. (Section 7)**

He: Table 36 report(s) correlation coefficients between the linear-effects and the additional terms (interactions). I would suggest including the canonical correlation between the set of linear-effects and the set of interactions. This would assess linearity beyond pairwise correlations. The considerations and justifications on finding final models (Section 7.2) were reasonable and well thought-out. Because some subjectivity was involved in the final model selections, it is hard for me to tell whether each detailed review and scrutiny reported in Section 7.3 is "optimal". On the other hand, there is no single model that is likely to be *the best*. In reality, it is generally the case that several models are (almost) equally good given the limited amount of data, and the final selection can be made with some subjectivity. The analyses given in Section 7.3 appear reasonable.

Lindhjem: The approach to modeling is well considered given the low emissions rates of these vehicles and the relatively small fuel effects, often below normal detection limits. It was apparent that the iterative process using not only novel statistical techniques (compared with other fuel effects evaluations), but an understanding of the testing limitations (low emission rates coupled with detection limits) was a necessary method to determine relevant fuel parameters to include in the evaluation.

I agree with the approach of limiting the number of statistically fit terms (especially the second order terms) to only those that assist in explaining the fuel effects. Perhaps the discussion of the magnitude of the residuals could further highlight the lack of impact of terms that have been dropped from the correlations such as in Figure 70, where a trend may exist, but is relatively small in magnitude.

The overall statistical approach is sound using several methods to discover and systematically eliminate terms. The approach to identify influential data used appropriate physical (investigate detection limits and other laboratory variables) and statistical (significance level, BIC, and residual evaluation) methods to eliminate or keep data.

West: This approach seems reasonable based on the discussion.

4. **Comment on the use of the Tobit regression (SAS PROC LIFEREG) for modeling datasets with large numbers of censored values. (Section 5.3)**

He: I think that the use of left-censored models and the Tobit regression is appropriate. My only question here is why the same approach is not taken for cases with light censoring.

I believe that the decision to use left-censored models in the study is appropriate. This allows linear models to remain valid for the data with left end points. Such practices have been used in the statistics and econometrics literature. There are, however, several minor issues to deal with.

(a) In Section 5.2, censored measurements were replaced by the minimum positive value measured for the emission and bag. This substitution has the potential to lower the variance estimate of the statistical models, unless censoring is taken into account in the variance estimates. If the error variance estimates are deflated, we would see higher “studentized residuals”, resulting in false positives in outlier detection. The issue needs to be carefully examined.

(b) In the analysis of Section 5.3.1, a distinction was made between light censoring and severe censoring. In the case of severe censoring, the Tobit regression was used in the data analysis. Otherwise, the censored values were substituted. I do not see good reasons for handling the two scenarios differently. Why not use the Tobit regression in all cases? The current practice would raise a question about the stability and sensitivity of the results if one more or one fewer data point is censored.

(c) In Section 5.3.2, both BIC (model selection criterion) and likelihood ratio tests were described and used. Although both approaches are valid and useful, they have different goals in mind. Model selection based on BIC is to choose models, treating all competing models equally. The chi-square tests have a null hypothesis in mind, where the null hypothesis refers to the smaller models in the present analysis. Test decisions are designed to protect the null hypothesis, so the competing models are not treated equally. When both approaches are used, one needs to be clear how they work together, and what are to be achieved. I do not imply that anything has gone wrong here, but this part of the analysis needs to be made clearer as to why one cannot simply use BIC.

West: This approach seems reasonable, although I would like to see “large” and “small” censoring levels explained further. Censoring of less than 5 values is considered small. Why? With over 900 datapoints, could the limit be set at a higher number? Please explain.

5. **Comment on the decision to independently model the linear terms and interactions between fuel blends as presented in the final results?**

He: Table 36 report(s) correlation coefficients between the linear-effects and the additional terms (interactions). I would suggest including the canonical correlation between the set of linear-effects and the set of interactions. This would assess linearity beyond pairwise correlations. The considerations and justifications on finding final models (Section 7.2) were reasonable and well thought-out. Because some subjectivity was involved in the final model selections, it is hard for me to tell whether each detailed review and scrutiny reported in Section 7.3 is “optimal”. On the other hand, there is no single model that is likely to be *the best*. In reality, it is generally the case that several models are (almost) equally good given the limited amount of data, and the final selection can be made with some subjectivity. The analyses given in Section 7.3 appear reasonable.

Lindhjem:

Statement (p. 9) “*The analysis involved ongoing and iterative interaction between statistical modeling and additional physical and chemical review of the data.*” Page 23 *“The final design is the result of an iterative process involving interactions between research goals, the feasibility of fuel blending, and experimental design.”*

Comment The term “interaction” should refer only to statistically fitted second order terms where fuel parameters are mixed, such as ZZ_{ea}, to represent the second order ethanol x aromatics term. The statement above appears to use “interaction” in a different context and so is confusing. EPA should also search all other uses of “interaction” to ensure that there is no confusion.

Statement (p. 13) “*Note that this generalization does not account for the effect of interactions between RVP and other properties, which are in some cases larger than the underlying linear effects.*”

Comment I see only two RVP interacting terms affecting only running THC (not NMOG or NMHC) and start CO, and this statement only appears to be true for the CO start emissions. I would suggest either stating this more plainly or striking the comment. Or does this line refer to fuel properties that are affected when RVP is modified? For example, reduced T50 or diluted aromatics occur when lighter compounds are added to increase RVP; then this statement would be true for the other fuel properties than just for RVP.

West: This approach seems reasonable based on the discussion.

6. Please comment on the methods used to select reduced models. (Section 5)

He: Table 36 report (s) correlation coefficients between the linear-effects and the additional terms (interactions). I would suggest including the canonical correlation between the set of linear-effects and the set of interactions. This would assess linearity beyond pairwise correlations. The considerations and justifications on finding final models (Section 7.2) were reasonable and well thought-out. Because some subjectivity was involved in the final model selections, it is hard for me to tell whether each detailed review and scrutiny reported in Section 7.3 is “optimal”. On the other hand, there is no single model that is likely to be *the best*. In reality, it is generally the case that several models are (almost) equally good given the limited amount of data, and the final selection can be made with some subjectivity. The analyses given in Section 7.3 appear reasonable.

West: This discussion was convincing, although I am not an expert in this area. I understand that reduced models have lower likelihood of the models describing the random error rather than the underlying fuel effects.

3.1.2 Other Specific Technical Comments

All of the reviewers provided specific technical comments in addition to their responses to the specific questions in the peer review charge.

He: I have some additional suggestions, some of which might be useful for future studies. First, if a similar study is planned in the future, it would be better to construct specific criteria for removing outliers prior to data collection. This would eliminate questions about biased interference in the data processing stage. Second, when imputation is used, more careful procedures should be in place to account for variability due to imputation. Multiple imputation is a common approach to take. Third, some sensitivity analysis using robust statistical methods can be performed to understand the effects of outlying/influential points and their handling on the final analysis. Most statistical techniques, including linear mixed models and the Tobit regression used in this study, are based on the assumption of Gaussian errors. Robust statistical methods can help us understand the impact of non-Gaussian errors.

Lindhjem:

[1] The fuel properties in testing matrix used in the evaluation are in two respects quite different from previous studies. The higher levels of ethanol, up to 20%, are beyond what any previous study has considered. Also, the lack of olefins evaluation data eliminated one fuel parameter that was found to have a significant effect on emissions in previous studies.

[2] One issue to be determined is how EPA or others will choose to extend the fuel property relationships developed in this report to the general fleet, such as in MOVES. Because vehicles will naturally age and may respond differently to fuel properties, these relationships may not continue to hold true. These higher emitting vehicles could potentially contribute to the emissions inventories out of proportion to their numbers. Care should be taken when extending the fuel effects to other vehicles, whether aged late model light-duty or heavy-duty gasoline powered vehicles.

[3] **Statement (p. 181)** *“[14] The typical hydrocarbon analyzer used for emission testing uses a flame ionization detector (FID), which is calibrated to accurately count carbon atoms that are bonded to hydrogen. Carbons bonded to oxygen, which occur in carbonyl and alcohol emissions from burning ethanol fuels, are not accurately counted by the FID, and thus emissions from ethanol fuels require additional characterization methods to properly quantify as NMOG or VOC.”*

Comment This statement is generally, but not strictly, accurate in that most FID units are calibrated on propane, and the hydrocarbon measurement assumes that all carbons in the sample respond the same as the carbon atoms in propane. Hydrocarbons do not necessarily respond identically as propane carbon atoms do, but carbons bound to oxygen respond at a rate order(s) of magnitude lower. The error for most hydrocarbons has been considered insignificant. The response to oxygenate carbons has been historically considered to be insignificant compared with other hydrocarbons in the sample, so carbonyl and alcohol compounds determined through alternative methods are added to the NMHC weight measure as the weight of single carbon aldehyde (formaldehyde) and alcohol (methanol). The different composition of hydrocarbons and any FID response to oxygen-bound carbon influence the NMHC measurement, but those influences are considered minor. With lower emission rates of these newer vehicles, these assumptions may not be as appropriate as for older vehicle designs with higher emission rates. In addition, does the 'typical hydrocarbon analyzer' differ from the $NMHC_{FID}$ measurement that needs to be correlated with NMHC before the statistically modeling proceeds?

[4] *"8.2.2.4 Vapor Pressure" (p. 182)*

Comment I found this section description (and others that rationalize the effect modeled) more of an unsatisfying hand waving exercise to justify what the data was telling. If indeed the older studies and this most recent evaluation are to be believed, then the modeled effect may be temperature dependent when at high temperature the feedback of vapors to the engine may affect the emissions opposite to that modeled here. The suggestion is to limit the speculation to the results of the analysis presented and note other references if the effect found needs to be justified.

West:

[1] In Section 7, mean residuals are plotted, presumably to demonstrate the quality of the models. It would be interesting to see range bars on the data points to show the range of individual residuals.

[2] (p. 9) This approach was followed for several reasons: (1) the candidate fuel effected identified for study were selected because we anticipated that they could be important for one or more emissions. ***Not clear. Candidate fuels? Or candidate fuel effects?***

[3] (p. 13) *Ethanol*: taken in isolation, the models indicate that increasing ethanol is associated with increases in all emissions, both for cold-start and hot-running emissions. The sole exception to the pattern is CO, for which the response to fuel properties appears to change between start and running. The effects are strongest for PM, NOx and NMOG, although presumably, the underlying physical processes could vary. ***Interestingly not consistent with V1 (ORNL/TM-2008/117) or V4 (ORNL/TM-2008/234). Increasing ethanol (in splash blends with certification gasoline) decreased NMHC and THC (FID_HC), while NMOG was relatively flat. Ethanol and acetaldehyde increased, of course. V1 used LA92 but V4 used FTp.***

[4] (p. 36) For these reasons, the decision was made not to replace zeros in the dilute bag dataset with integrated continuous measurements. ***Agree with this decision.***

[5] (p. 50) At the outset, it is helpful to get an overview of the raw results, sorted by vehicle and fuel, which gives an initial impression of variability among vehicles and fuels, as well as within vehicles. ***Agreed. Also would be helpful to show variability of individual vehicles on individual fuels.***

[6] (p. 51) In the plot for etOH×T50 (Figure 18), the view seems to indicate an upward trend from E0 through E20, but with some downward curvature above E10. ***Show all data? How much scatter is there in NOx for these cases?***

[7] (p. 52) At first glance, the trend appears to "zig-zag," from low to high. ***Test to test variation of Bag 1 NOx in V1 study (LA92 cycle) ranged from 10% to over 100% (same veh, same fuel). In V4 program (FTP, not LA92), range of weighted composite NOx approached 100% in some cases. How do you ensure that coincidental test-to-test variation is not erroneously attributed to a fuel effect?***

[8] (p. 76) The Linear Effects plot for ethanol shows some mixed results (Figure 33), but with an apparent increase from 0% to 10% ethanol, followed by a leveling or decline at higher ethanol levels. ***Decline would be expectation. Increase from E0 to E10 is odd.***

[9] (p. 93) In this initial step, censored measurements were replaced with the minimum positive value measured for the emission and bag. ***Minimum across all vehicles and fuels?***

[10] (p. 96) For minimal levels of censoring, defined as five or fewer censored measurements ($n_{\text{censored}} \leq 5$), we elected to substitute the minimum positive measured value for the missing measurements. After substitution we fit mixed models as described above. ***N<5 seems reasonable if Total number of samples is large. Suggest noting Ntotal here. Why 5? That is, why not 4 or 6 or 10? Is 5 arbitrary or is there a citable reference to why 5?***

[11] (p. 100) In this case, the BIC declines steadily as terms are removed, indicating an improvement in fit for each successive reduced model. ***Declining steadily by 0.1 to 0.2% does not seem significant or important. Is the improvement cited below truly due to the simpler model? Perhaps, but is the BIC really an accurate indicator?***

[12] (p. 101) Table 20. Model Fitting History for PM, Bag 1 (FM9 selected as best-fit model). ***Is the sensitivity of BIC such that there is a significant difference between 2862 and 2867? Six significant figures and 0.1% change in BIC does not seem important. If this is important or significant, it should be explained. Perhaps the important point is that BIC is NOT INCREASING as the model is simplified, thus justifying the simpler model?***

[13] (p. 120) In this program, dilution air was HEPA-filtered and presumed to be free of PM, so there was no background filter sample collected for later subtraction as is typical with other emissions. ***Why no tunnel blanks to prove this assumption?***

[14] (p. 120) Discussion with EPA staff experienced with PM measurement suggests that for the data as collected in this program a variability of ±1 µg should be applied to all filter weights. Considering that the net PM result is calculated by subtracting two filter weights (average dirty minus average clean), it should be understood to have a variability range of ±2 µg, as the measurement error applies to both weights. Therefore, a net weight gain of 10 µg would have a relative error of 20% associated with it, a figure of the same order of magnitude as the fuel effects this program attempts to capture. ***Need to cite reference(s) or present data to establish this level of error. For example 2005-01-0193.***

[15] (p. 149) Interaction plots for selected terms are shown in **Figure** . ***Y axis title is "mean measurement." Isn't this a modeled result, not a measured result?***

[16] (p. 149) Another way of viewing the interactions is to average and plot the residuals of the linear effects model. ***Compare modeled results to measured results. Mean is good, but would***

also be nice to see the range of residuals. Mean can be close to zero, but how much variation is there? (average of + 20 and -20 is zero).

[17] (p. 161) Figure 69. NOx (Bag 1): Mean Residuals for the Linear Effects Model, vs. Target Fuel Properties for four pairs of terms: (a) Ethanol × Aromatics, (b) Aromatics ethanol, (c) Aromatics × T90, (d) T90×Aromatics, (e) RVP × T90, (f) T90 × RVP, (g) ethanol × T50, (h) T50 × etOH. ***Same comment as above. Mean residuals look good. What is range of individual residuals? Perhaps add error bars to show min and max or perhaps interquartile range of residuals?***

[18] (p. 178) This question does not arise in the context of model application, but rather with respect to model validation, in that the datasets available for validation include data that represent emissions from pre Tier-2 vehicles. ***Such as the DOE V1 and/or V4 datasets? Have the models been compared/validated against V1 or V4 data? Note that V1 used splash blends and ran the LA92 cycle. V4 tests also used splash blends but ran the FTP.***

[19] (pp. 180-81) Thus, despite much lower overall emission levels that have been achieved in Tier 2 vehicles through improved fuel control and catalyst efficiency, the effect of ethanol on combustion **and aftertreatment(?)** appears to persist in certain modes of operation such as cold-starts and transients during warmed-up operation. ***Effect of ethanol on NOx emissions may be related to exhaust stoichiometry and catalyst efficiency more so than (or in addition to) changes in engine-out NOx.***

[20] (p. 181) (Should we address effect on THC/NMHC as well?) ***Yes. Note that NMOG in V4 was not affected by ethanol, but FID_HC and NMHC decreased (with splash blends).***

[21] (p. 181) In the present study NMOG decreased with **decreasing** aromatics content, in agreement with earlier studies. ***Or increased with increasing aromatics…***

[22] (p. 199) With respect to censoring, the following rule was applied. If the number of censored measurements was ≤ 5, we substituted the smallest measured positive value for the missing values, and proceeded with model fitting, using a mixed-model approach. ***Why 5? Why not 4 or 6 or 10? Explain.***

[23] (p. 199) However, if the number of censored measurements was > 5, we fit a model using Tobit regression (i.e., "censored normal regression"), an established technique for analysis of left-censored datasets. ***Ditto. What is significance of 5?***

[24] (p. 203) Based on these results, the reduced model FM7 was selected as the best fit. ***As noted in previous BIC discussion, do small variations in BIC truly indicate a difference in fit? FM5, FM6, and FM7 all have similar BIC.***

3.2 General Comments

The reviewers provided general comments on the *EPAct Study Analysis*. Among these general comments were evaluations of the report's strengths, suggestions for improving and strengthening certain of its elements, and queries for further information.

He: Overall, I found the study well designed and carefully analyzed with generally accepted modern statistical tools. My comments will focus on the appropriateness of the data processing and statistical techniques described in the report, with the purpose of improving the EPAct Study Analysis.

Lindhjem:

[1] The report . . . presents a well-documented approach to estimating the effect that gasoline fuel properties have on late model vehicle emissions. In general, the approach to evaluating the data is equal to or more robust than previous efforts, such as the complex and predictive models or other fuel effects studies.

[2] I was unable to find the earlier reports "EPAct/V2/E-89" referenced in the document, presumably "Assessing the Effect of Five Gasoline Properties on Exhaust Emissions from Light-Duty Vehicles certified to Tier-2 Standards: Part I - Study Design and Execution (EPAct/V2/E-89Test Program Final Report)". This report likely describes in better detail why the fuel blending program could not produce an orthogonal fuel matrix. Likewise, questions about the measurement, vehicle conditioning, and other issues are not addressed in Part II. This makes it difficult to assess the relevance of the modeling.

West:

[1] The report is long and contains tremendous detail regarding the statistics and modeling approaches used in analysis of the EPAct/V2 data from the SwRI program. Considering its length and the technical detail, the report is well written and well organized.

[2] This reviewer has over 20 years experience in engine, emissions, and vehicle testing, but limited experience with many of the statistical methods and modeling approaches described. To the extent possible in the short time available, methods and terms were researched and explored while reviewing the EPA document. Explanations and approaches appear largely reasonable. Some specific questions are noted in comments in the report. For example, the Bayesian Information Criterion (BIC) is used to determine goodness of fit. In one example, the BIC for the full model starts out at >2900 and decreases by 0.1-0.2% for subsequent reduced models. The discussion refers to "a steady decrease" in BIC. The reviewer suggests the significance of such a small change in BIC be discussed further. While the reduction in BIC appears *insignificant*, the fact that BIC is *not increasing* is perhaps the justification to use the reduced model (sec. 5.3.2).

[3] The authors appear to have taken great care to ensure that models are as representative as possible, and to avoid overly complex models or overfitting. The handling of "nondetects" appears reasonable. There is rigorous treatment and discussion surrounding background measurements, analyzer drift, limits of quantitation, censoring of data, etc. Nonetheless, the complexity of the problem makes it difficult to see whether all of the objectives were achieved or rather that some of the apparent results are any more than artifacts of an intricate math problem. As an example, all emissions data are shown by vehicle and include all tests on all fuels, to show the range of measurements for a given pollutant, by vehicle. These are very informative charts. However, no data are shown to demonstrate the test-to-test repeatability for a given vehicle with a given fuel. Several figures (such as Figures 13-21) show averages, but would be more informative if range bars were shown to indicate max and min or perhaps interquartile range, or perhaps scatterplots with all data points.

[4] Authors discuss taking care to not overfit or model random scatter (which is good), in fact stating that the Bag 3 models may be more vulnerable to measurement error due to the extremely low emissions, and opting not to report model coefficients for Bag 3 results. The reviewer agrees with and commends this decision. Handling of extremely low measurements is difficult, and the authors discuss this issue extensively and convincingly. However, when it comes to the PM emissions measurement, there could be additional discussion and justification. For instance, one issue that deserves additional discussion is the lack of tunnel blanks for background PM. Data from a few example tunnel blanks could provide convincing evidence of the zero background assumption. Furthermore, control data should be shown, or at least cited, to support the assumed level of error (± 1 µg) in the PM filter weight measurements. The authors mention discussions with EPA staff as the basis for the assumed level of error. As the authors know, PM measurement is very sensitive to measurement precision and accuracy, with temperature, humidity, buoyancy effects, static charge and other factors greatly influencing results. In SAE 2005-01-0193, the authors detail exhaustive measures to attain an accuracy of better than ± 1 µg in their filter weights. Please provide evidence of the stated PM measurement error.

[5] Models should be validated against independent datasets, as discussed in section 8.

[6] (p.8) The program was conducted in three phases. Phases 1 and 2 were pilot efforts involving measurements on 19 light-duty cars and trucks on three fuels, at two temperatures. This work was completed at Southwest Research Institute between September 2007 and January 2009. ***Have these results been published?***

[7] (p. 16) This report describes the analysis of the dataset collected in Phase 3 of the EPAct/V2/E-89 program, conducted at Southwest Research Institute in San Antonio, Texas. A separate report describing the program design and data collection activities is available, but an overview is provided below. ***This report does not appear to be publicly available at this time (January 2012)***

[8] (p. 18) An initial sample of 19 test vehicles was chosen with the intent of representing **~~of~~ the** latest-technology light duty vehicles being sold at the time the program was being launched (model year 2008). In terms of regulatory standards, the test sample was to conform on average to Tier 2 Bin 5 exhaust levels and employ a variety of emission control technologies, to be achieved by including a range of vehicle sizes and manufacturers. ***I recall EPA staff indicating (c. 2007) that the list of vehicles was a projection of future technology and/or engine families***

[9] (p. 20) After some consideration, study participants agreed to rely on the aggregate data, while applying appropriate techniques **~~do~~ to** address the resulting "censoring" of the data at low end of the range of values. ***Who are the "study participants?"***

[10] (p. 31) As one-stage standardization did not neutralize correlations among model terms, we applied a second stage of standardization to the 2nd order terms. ***This report does not appear to be publicly available at this time (Jan 2012)***

[11] (p. 32) Table 8 shows that the combination of one- and two-stage standardization neutralizes the remaining correlations, with the exception that between the etOH and T50 linear effects, as previously described. ***Unclear "that between the etOH and T50 linear effects (what)"***

[12] (pp. 36-37) Based on strong correlations between these species, we developed statistical models to impute NMOG and NMHC from corresponding $NMHC_{FID}$ measurements. ***See ORNL/TM-2011-461***

[13] (p. 51) Trends for individual vehicles show a general increase in NOx with increasing ethanol, with some exceptions. ***Consistent with prior studies ORNL/TM-2011/234; SAE 2009-01-2723***

[14] (p. 91) We assume that a very small but positive measurement existed but was not captured and quantified. Assigning a value of zero to these observations is an example of a common approach to censoring of observations, known as "substitution." In this approach, a small but fixed quantity is substituted for the censored observations. Values used for substitution include zero, as mentioned, or small but positive quantities such as the smallest observation, a multiple of the smallest observation, the limit of quantitation (LOQ) or half the limit of quantitation (LOQ/2). The degree of censoring varied widely by emission and bag, as shown in Table 2. At different stages of the analysis, we addressed censoring in different ways. ***These various approaches all make sense. Suggest adding couple of sentences about the various bags of the LA92. Bag 1 is cold with majority of emissions, hence less censored values. Bag 2 is hot, but includes some open-loop operation due to hard acceleration. Bag 3 is hot start bag with very low emissions, hence majority of censored values.***

[15] (p. 92) Table 15. Numbers of Censored Measurements, by Emission and Bag. ***Suggest showing total number of tests (in title or footnote).***

[16] (p. 94) Table 17. Counts of Influential Measurements, by Emission and Bag (with "influential" defined as having a studentized-deleted residual ≥ 3.5 or ≤ -3.5). ***Suggest including total number of measurements or observations***

[17] (p. 95) An additional measurement in Bag 3 (run 6281) was removed, even though it was not flagged as influential. ***Explain further?***

[18] (p. 95) The full sets of terms in the optimized design include terms anticipated to be meaningful for any of the emissions to be measured. However, it was not anticipated that all the terms included would necessarily be meaningful for all emissions in all bags. A closely related goal is to develop models that would be, to the extent possible, explicable in terms of knowledge of the relevant physical and chemical processes. Parsimonious models are preferred over full models for this purpose, as their simpler structure makes their behavior easier to assess and explain. Finally, with respect to explicability, it is much preferred to minimize the potential for overfitting, which could reduce the generality of models selected for prediction. To guide the process, we adopted several assumptions, described below. ***Model describes the random error rather than the desired underlying relationship. Good discussion.***

[19] (p. 107) The approach to analysis of censored measurements, as described in 5.3.2, was also adopted based on guidance from the author of the DOE research. ***No public documents can be found on the EPAct/V2 study.***

[20] (p. 116) If this program were simply trying to quantify the magnitude of NOx emissions from such vehicles, this level of error may be acceptable. However, since we are looking for meaningful differences in emissions between fuels, this large relative error is particularly problematic. ***Key point. Meaningful differences.***

[21] (p. 117) This suggests it likely has higher measurement noise than data from the other vehicles, and thus many measurements may not be reliably distinguishable from background levels. ***Good***

[22] (p. 149) It is interesting to note that while (d) and (f) appear somewhat similar visually, the model considers the aromxT90 interaction highly significant but the RVP×T90 interaction insignificant. ***Explain.***

[23] (p. 150) < insert physical interpretation of these interactions here?> ***Yes***

[24] (p. 165) One is that the magnitudes of corresponding coefficients are generally larger for Bag 1 than for Bag 2 emissions, suggesting that the effects of fuel properties are more pronounced for "cold start" than for "hot running" emissions. ***As expected!***

[25] (p. 169) The reasons for these differences are not apparent, but it is clear that the relations among NOx, etOH and T50 are complex ***Or nonexistent?***

[26] (p. 169) It may be appropriate to consider whether the Bag 3 results may be more vulnerable to measurement error attributable to low sample measurements relative to background, given the issues with measurement discussed in 6.1.1 (page 116). ***Yes***

[27] (p. 183) Consider showing quantitative parameter changes and percent change results for the models?) ***Good idea***

[28] (p. 211) Detailed results, including models fit, fitting histories, coefficients and tests of effect are presented in Appendices Q.3-W.3 for Bag 1 models, and Q.4 - W.4 for Bag 2 models. ***Not available for review***

3.3 Editorial Comments

Two of the reviewers undertook a thorough editing of the report, providing significant editorial comments. These reviewers noted typographical and formatting errors, incorrect word choice, and omissions, including missing references.

Lindhjem: There are numerous editorial corrections that need to be made, but most have already been noted in the document itself or are obvious from WORD program review. What follows are specific suggestions either for improving the flow or for feedback that would not otherwise be considered during normal editing. The original line is provided in *"quoted italics"* and the suggestions, questions, or comments appear in suggested edits (strikethrough or red) or flat text.

Statement (p. 9) "*The models reported in this section are as ~~parsimonious~~* ***concise*** *as the data and subject-matter knowledge allow.*"

Statement (p.13) "*Ethanol: taken in isolation, the models indicate that increasing ethanol is associated with increases in all emissions, both for cold-start and hot-running emissions. The sole exception to the pattern is CO, for which the response to* ***added ethanol is lower CO emissions during start but inconclusive for running conditions.*** ~~**fuel properties appears to change** ***between start and running.***~~"

p. 22 [5] *“This parameter was measured as DVPE, but **f**or simplicity and consistency, we will refer to it as “RVP.””*

West: Several recommended edits and comments are provided in the marked up report, to correct typos or improve clarity. As an example, many figures could use larger fonts before final publication. Several cited references (especially Part I of the same report and the appendices) appear unavailable at present, and there are a few missing figures and references. But overall it is a very good draft.

(p. 8) An initial sample of 19 test vehicles was chosen with the intent of representing **~~of~~ the** latest-technology light duty vehicles being sold at the time the program was being launched (model year 2008).

(p.9) Speciation also allowed independent analyses of **selected** toxics including acetaldehyde, formaldehyde, acrolein, benzene and 1,3-butadiene.

(p. 10) In generation of emissions, the effects of different fuel properties are not separable, in that it is difficult to modify one **property** without affecting one or more of the others.

(p. 10) However, the coefficients for different fuel properties can be directly compared, allowing assessment of the relative importance of the effects of the fuel properties on the emissions **constituent being** modeled.

(p. 13) Cold-start **(Bag 1)** CO and hot-running THC both have small decreases in emissions with increasing aromatics. In terms of magnitude, the pattern is similar to ethanol, with PM, NOx and NMOG showing the strongest effects.

(p. 16) Since data on Tier 2 vehicles are critical to understanding the impact of fuel property changes on the onroad vehicle fleet as increasing volumes of biofuels **are introduced**, EPA entered a partnership with DOE and CRC **to** undertake the largest fuels research program conducted since the Auto/Oil program in the early 1990s. This program is aimed specifically at understanding the effects of fuel property changes on regulated and **selected** unregulated exhaust emissions from later technology Tier 2 vehicles.

(p. 17) These five parameters were selected based on previous studies on older vehicles as having potential to affect exhaust emissions [cite]. ***Missing reference.***

(p. 17) The parameter ranges to be covered for T50, T90, aromatic content, and RVP were selected to represent the range of in-use fuels based on a review of the Alliance of Automobile Manufacturers’ 2006 North American Fuel Survey [cite?]. ***Missing reference.***

(p. 17) Test fuel parameter ranges were originally drafted to span roughly the 5th to 95th percentiles of survey results **~~for~~** in U.S. gasoline, though some test fuel parameters were adjusted after the actual blending process began.

(p. 18) An initial sample of 19 test vehicles was chosen with the intent of representing **~~of~~ the** latest-technology light duty vehicles being sold at the time the program was being launched (model year 2008).

(p. 20) After some consideration, study participants agreed to rely on the aggregate data, while applying appropriate techniques ~~do~~ **to** address the resulting "censoring" of the data at low end of the range of values.

(p. 21) The methods used were very similar to those used for the modeling **of** the other emissions, with modifications to address issues of study design and measurement specific to these compounds.

(p. 22) The design and implementation of the study, including the aspects of fuel blending, measurement methods and logistics are described in a separate report4.**Error! Bookmark not defined.** ***Missing reference.***

(p. 22) It is well known for fuel properties **to** be moderately to strongly correlated.

(p. 22) [1] This parameter was measured as DVPE, but for simplicity and consistency, we will refer to it as "RVP."

(p. 25) Measurement methods are discussed in detail in the testing report4. ***Missing reference.***

(p. 27) In addition to correlations among the linear effects and interactions, and correlations among interactions, we can see one fairly strong correlation~~s~~ among the linear effects, specifically, between etOH and T50 (R = -0.57).

(p.33) In this **constant volume sampling** system, the vehicle exhaust is mixed with a large amount of filtered dilution air, and a small portion of this stream is continuously withdrawn to fill a sealed bag over the course of a test cycle.

(p. 33) Its primary disadvantage is that the **overall** dilution ratio of background air to exhaust must be fixed for an entire test, and is set relatively high to avoid condensation of water vapor within the system during periods of high exhaust flow.

(p. 35) [1] This assessment of the situation ignores the possibility that the vehicle actually consumes or destroys a given pollutant species during parts of the test cycle, resulting in periods of "negative emissions", such that the average emission level over a test is truly zero. While situation**s** may occur over a limited period for some emissions in a highly polluted environment, e.g., PM or NMHC in congested traffic, it is highly unlikely in an emission test cell.

(p. 36) Rather, these measurements ~~to~~ **were** treated as "censored."

(p. 36) The alternate measure "NMHC as measured by FID" ($NMHC_{FID}$), was collected for the entire dataset, and it very tight**e**ly correlated with both NMOG and "true" NMHC.

(p. 37) On this basis, we fit linear models for NMOG and NMHC in terms **of** $NMHC_{FID}$.

(p. 37) Scatterplots of NMOG vs. $NMHC_{FID}$ for Bag 2 are presented in ***Formatting.***

(p. 38) For the bag 2 and 3 models, the counterpart to Equation 8 is Equation 11, which simplifies for blends other **th**an E0 similarly to Equation 8, except that the slope term is always as in Equation 9.

(p. 51) A linear-effects plot for ethanol is shown in Figure 13, which suggests that an ethanol effect is visible when the data ~~is~~ **are** averaged across **the** other four fuel properties.

(p. 52) It is necessary to go a step further, and look at "interaction" or "conditional effects" plots, starting with the interaction of ethanol and T50, which deserves special attention **because**….

(p. 64) As in Bag 1, variability for most individual vehicles spans one third to half an order of magnitude. ***Missing reference?***

(p. 64) The effects for T50 and T90 are similar in that the trends across the five T50 levels and three T90 levels are similar, with no overall trend apparent, for the same reasons noted above for Bag 1 (Figure 26, Figure 27). ***Missing reference?***

(p. 76) The variability within vehicles is about 1-1.25 orders of magnitude. ***Missing reference?***

(p. 77) The Linear Effects plot for T90 is clearly suggestive of **an** overall positive effect, when considering all vehicles (Figure 37).

(p. 77) The left-hand point represents the same fuel as **the** left-hand point in the green trend in the previous plot.

(p. 77) In contrast, the plots for ethanol and aromatics are suggestive of a positive or "reinforcement" interaction. In the etOH × arom view (Figure 40), the trend for the higher aromatics level (green) appears steeper than for the lower aromatics level (black). Similarly, in the arom × etOH view (***Font.***

(p. 77) In Figure 41.5, the etOH × T90 plot does not appear to suggest interaction, if we discount the green trend (T90=325°) as representing only two fuels. ***Missing figures***

(p. 99) At each step, we tested the goodness-of-fit of each reduced model against that of the full **model** using a likelihood-ratio test.

(p. 102) All models based on the 16-parameter full model, as shown in Equation 14 **are**…..

(p. 107) Thus, in running these models, censored values were replaced ~~in~~ **with** the minimum positive measured value in each bag for each emission.

(p. 116) This **observation** led to a closer examination of measurement error in the dataset.

(p. 117) This **condition** would be expected to give a zero result as discussed in Section 3.1.

(p.118) The data for the Sienna, though similar in their range of sample and background measurements, ~~was~~ **were** not found to be exceptionally influential to model-fitting and therefore ~~was~~ **were** not removed from the dataset.

(p. 120) After the test is conducted, the filters are removed, **placed** back into the clean containers, and returned to the clean room.

(p. 124) Figure 5**6**~~4~~ and Figure 5**7**~~5~~ show net measurements as a percentage of sample for bags 2 and 3, as an attempt to understand the magnitude of the nets relative to the measurement error where nets are small.

(p. 124) Two vehicles (Odyssey, Sienna) have the majority **of** points below those of other vehicles.

(p. 124) Plots of ambient and sample for bags 1-**3**~~2~~ are shown in Figure 58, Figure 59, and Figure 60, followed by examination of nets as percentage of sample for bags 2 and 3 in Figure 61 and Figure 62.

(p. 138) At this point it is important to note that the results reported below differ from those reported above in 5.3, as well as from those reported for the DOE analysis8. ***Missing reference***

(p. 151) Figure 65. ln(CO) (Bag 1): Two-way Conditional Effects Plots for Three Interactions and one Quadratic term, viewed with respect to Both fuel Parameters : (a) Ethanol×Aromatics, (b) Aromatics×Ethanol, (c) Aromatics×T90, (d) T90×Aromatics, (e) RVP×T90, (f) T90×RVP, (g) etOH×etOH. ***Typo in fig c and e. T90 blue should be 300***

(p. 153) Figure 66. CO (Bag 1): Mean Residuals for the Linear Effects Model, vs. Target Fuel Properties for three Interactions: (a) Ethanol × Aromatics, (b) Aromatics × T90, and (c) RVP × T90, and the quadratic term etOH×etOH. ***Blue T90=300***

(p. 158) Another possibility is that the effect **of** T50 in this presentation is masked by variation in RVP and T90 levels across these fuels.

(p. 164) The relations among the ethanol, aromatics and RVP coefficients are similar to Bag 1 and to each other, except that the aromatics coefficient is slightly higher than the ethanol coefficient for NMOG, while ethanol is more important ~~to~~ than aromatics for NMHC.

(p. 164) As in Bag 1, the etOH**x**arom and etOH**x**T50 interactions are reinforcements, whereas the etOHxRVP interaction is an interference (etOH positive, RVP negative, interaction negative).

(p. 169) When starting with the 16-term model, the reduced model ~~retains~~ includes RVP and T90 linear terms (both insignificant), plus two interactions not included in the design model: arom×T90 and RVP×T90.

(p. 176) *Ethanol*: taken in isolation, the models indicate that increasing ethanol is associated with increase **in** all emissions, both in bags 1 and 2.

(p. 180) The results of the present study are consistent, showing an increase in NOx emissions with an increase in ethanol level, regardless of whether it is from a statistical analysis of an orthogonal change **in** ethanol alone or when changes in other fuel properties typical of splash or match blending are included.

(pp. 180-81) Thus, despite much lower overall emission levels that have been achieved in Tier 2 vehicles through improved fuel control and catalyst efficiency, the effect of ethanol on combustion **and aftertreatment(?)** appears to persist in certain modes of operation such as cold-starts and transients during warmed-up operation.

(p. 181) In the present study NMOG decreased with **decreasing** aromatics content, in agreement with earlier studies.

(p. 182) However, during cold start, the new data suggests the direction of **the** effect is dependent on ethanol content.

(p. 182) The present study shows little or no effect of T90 on NOx, which is reasonable considering that NOx production and control are largely about heat release, and T90 represents a smaller amount of combustible material at cold-start temperatures, and thus less energy, than T50. ***Not clear, please review/reword.***

(p. 184) More variation is seen in T50 and RVP, which span somewhat wider ranges of 160-195 °F and 7.3-13.0 ~~lb~~ **psi**, respectively.

(p. 184) Use of the non-standardized models is much more straightforward, in that they allow input of fuel properties in their original units, e.g., % for ethanol and aromatics, ~~lb~~ **psi** for RVP and °F for T50 and T90.

(p. 185) The processes and methods for hydrocarbon speciation are described in greater detail in the testing report4. ***Missing reference***

(p. 187) The "*G*-efficiency" for the full design was estimated at 51.6% for the eleven design parameters, as previously described in 2.1 ~~2.1~~ above (page 22).

(p. 187) As shown in the table, the design efficiency drops sharply **with** inclusion of 2nd order terms, to less than 5% for designs four and five.

(p. 190) In addition to estimating the sample measurements, as shown above, we estimated two variances, the first being the variance of the ~~variance of the~~ 5-day moving average of the media blanks ($\hat{\sigma}_k^2$), and the second being a variance of random errors ($\hat{\sigma}_\varepsilon^2$).

(p. 193) The first two plots (Figure 8**1**~~2~~) show acetaldehyde vs. ethanol, by T50 level, in linear and logarithmic space.

(p. 202) If the *p*-value for the test-against-previous is greater than the critical value, the null hypothesis of no significant differen**ce**~~t~~ in fit between the reference and nested models is retained, and the nested model is retained as the current best fit.

(p. 208) Finally, to illustrate the results of the jackknife replication procedure, Figure 83 shows cumulative distributions of coefficients for each jackknife replicate for the five ~~linear~~ linear **e**~~E~~ffects.

(p. 208) The distributions for aromatics (Z_a),T50 (Z_5), and T90 (Z_9) are similar in that they show noticeable lengthening in the lower tail, suggesting that 2-3 vehicle**s** may be influential in decreasing the values of the coefficients.

(p. 214) 10 References: ***Refs 4 and 8 cannot be found***

4. References

U.S. EPA (2011). *Assessing the Effect of Five Gasoline Properties on Exhaust Emissions from Light-Duty Vehicles certified to Tier-2 Standards (EPAct/V2/E-89: Phase 3) – Part II: Data Analysis and Model Development*.

Appendix A: Resumes of Peer Reviewers

CURRICULUM VITAE

NAME: He, Xuming

HOME ADDRESS: 6577 Heron Ct., Ann Arbor, MI 48103

Phone: (217)-417-2880

E-mail: he_uiuc@yahoo.com

EDUCATION BACKGROUND:

Ph.D. in Statistics, 1989, University of Illinois at Urbana-Champaign, USA M.S. in Mathematics, 1988, University of Illinois at Urbana-Champaign, USA B.S. in Mathematics, 1984, Fudan University, Shanghai, China.

PRESENT AND FORMER POSITIONS HELD:

- Professor (2011 - present), H. C. Carver Professor of Statistics, Department of Statistics, University of Michigan
- Professor (2002-present), Professor and Jerry and Ann Nerad Professorial Scholar at the College of LAS (2005-2008), Affiliated Professor at Department of Computer Sciences (2008-), Associate Professor (1996-2002), Assistant Professor (1993-1996), Department of Statistics, University of Illinois at Urbana-Champaign
- Program Director of Statistics, National Science Foundation (2003-2005)
- Director, Illinois Statistics Office (statistical consulting office), University of Illinois at Urbana- Champaign (2000-2003)
- Director of the Scientific Committee, Center for Statistical Sciences, Chinese Academy of Science, Beijing (2006-)
- Honorary Professor of Statistics at the School of Management, Fudan University (2008-)
- Visiting Changjiang Professor at Northeast Normal University under Chung Kong Scholarship (2008-2010)
- Statistical Consultant, Argonne National Laboratory (1995-2000)
- Honorary Professor, Department of Statistics and Actuarial Science, The University of Hong Kong (2011-)
- Lecturer, Department of Mathematics, National University of Singapore (1989-1993)

RESEARCH INTERESTS:

Theory and methodology of statistical inference, with applications to data analysis from a wide range of areas

TEACHING EXPERIENCE:

Have taught statistics courses at the undergraduate and graduate levels since 1989

Ph.D STUDENTS:

Have directed doctoral dissertations in statistics and biostatistics of twenty students

EDITORIAL SERVICES:

- Co-Editor-Elect (2011) of *Journal of the American Statistical Association, Theory & Methods*.
- Editor of the IMS Bulletin (2007-2010); Associate Editor of *Journal of the American Statistical Association* (2004-2010), *Journal of the Royal Statistical Society: Series B* (2010-present), *Statistica Sinica* (1996-2002, 2005-present), and *The Annals of Statistics* (2004-present); Co- Editor of *Sankhya* (2004-2007); Co-Editor of *Statistics and its Interface* (2007- 2010); Associate Editor of *Statistics and Probability Letters* (1994-2010); Editor of *Journal of Multivariate Analysis* (1998-2004); Guest Editor for a special issue of *Science in China*, *Series A* (2009) and *Computational Statistics & Data Analysis* (2010).
- Referee for major statistics journals including *The American Statistician*, *The Annals of Statistics*, *Journal of the American Statistical Association, Biometrika*, *Biometrics, Bioinformatics, Journal of Multivariate Statistics, Journal of Statistical Planning and Inferences, Journal of Nonparametric Statistics, Journal of the Royal Statistical Society, Journal of Computational and Graphical Statistics, Annals of Institute of Statistical Mathematics, Neural Networks, Econometric Theory*, and several IEEE journals.

OTHER PROFESSIONAL SERVICES AND RECOGNITIONS:

- Elected Fellow of the American Association for the Advancement of Science (AAAS);
- Elected Fellow, American Statistical Association (ASA);
- Elected Fellow, Institute of Mathematical Statistics (IMS);
- Medallion Lecturer, 2007 Joint Statistical Meetings;
- Special Invited and Keynote Speaker at the Korean Statistical Society Fall Meeting, 2008;
- Distinguished Lecturer at the 1st IMS Asia Pacific-Rim Meeting in Seoul, Korea, 2009;
- Plenary speaker, the 2010 International Conference on Robust Statistics, June 28-July 2, 2010, Prague, Czech Republic;
- Chair-designate of the ASA Committee on Meetings (2011); Served in the ASA Committee on Scientific Freedom and Human Rights (1998-2000), ASA Committee on Noether Awards (2005- 2011), Search Committee for the ASA Executive Director (2007), ASA Committee on Federally Funded Research (2006-2010), and ASA Committee on Meetings (2008-2010);
- President of the International Chinese Statistical Association (ICSA) (2010); Elected board member of ICSA (2003-2004); Chair of the ICSA Nomination Committee (2005, 2006); Member of the ICSA ad hoc Committee on a new Applied Statistics Journal (2006-2007);
- IMS Council Member (2004-2007 by election, 2007-2010 ex officio); Editor of *IMS Bulletin* (2007-2010); Chair of the IMS Nomination Committee (2005-2006); Chair of the IMS Travel Award Committee (2004-2006); Chair of IMS ad hoc committee on IMS-China (2007-2008);
- Scientific Committee member of the International Conference on Robust Statistics (2004-2011);
- Chair of the Program Committee of the 58th World Statistics Congress of the International Statistical Institute (*ISI*) to be held in Hong Kong 2013;
- Program Chair for 2010 Joint Statistical Meetings in Vancouver; Organizer of Invited Sessions at the 1997, 2000, 2007, and 2011 Joint Statistical Meetings;
- Chair of the International Program Committee, Conference on Frontiers of Statistics: Biostatistics and Bioinformatics, July 06, China;
- Co-Chair of the Scientific Committee of the IMS-China International Conference, 2008;

- Review Committee for the Institute of Mathematics and Its Applications (IMA), 2009;
- Member of Scientific Committee, International Conference on Statistics and Society, July 10-12, 2010, Beijing;
- Co-Chair for the track *Quantile regression and semiparametric methods* at the 3rd International Conference of the ERCIM WG on COMPUTING & STATISTICS, December 10-12, 2010, London.

RESEARCH GRANTS, AWARDS AND SPONSORS

- PI on the National Science Foundation Award DMS-1007396 on *Efficient Modeling in Quantile Regression* (08/10 - 08/13)
- PI on National Institute of Heath Grant 1 R21 CA129671 01 A1 on *Nonparametric analysis of reverse-phase protein lysate array data* (with Dr. J. Hu at M D Anderson Cancer Center) (08/09 - 07/11)
- PI on National Institute of Heath Grant 1 R01 GM080503 01 A1 on *Low-rank Approximation to Probe-level Data with Application to Exon Tiling Arrays* (with Dr. J. Hu at M D Anderson Cancer Center) (07/08 - 05/12)
- PI on the National Science Foundation Award DMS-0604229 on *Inferential Methods for Quantile Regression* (08/06 - 08/10)
- Co-PI on the National Science Foundation Award DMS-0724752 on *Statistical evaluation of model-based uncertainties leading to improved climate change projections* (09/07 - 08/11)
- PI on the National Natural Science Foundation of China Grant 10828102 on *Theory and applications of nonparametric methods* (with Dr. Ning-Zhong Shi at Northeast Normal University, China) (01/09 - 12/10)
- Statistician on the National Institute of Health Grant R21HL090455-01 on *Muscle quality, exercise and weight loss in older women: A quantitative MRI study* (2007-2009) (PI: John Georgiadis)
- Senior personnel on the National Science Foundation Award DMS-0800631 on *Statistical Approaches to Integration of Mass Spectral and Genomic Data of Yeast Histone Modifications* (06/08 - 06/12) (PI: Ping Ma)
- PI on the National Science Foundation Award DMS-0630950 on *A Virtual Center to Promote Collaboration between the US- and China-based Researchers in Statistical Science* (08/06 - 07/09)
- Co-PI of the National Institute of Health Grant R01 DC005603 *Internet Evaluation of Oropharyngeal Swallowing Function* (03/04-02/09) (PI: Adrienne Perlman)
- PI of the National Science Foundation Award DMS-0102411 on *Quantile Regression* (07/01-07/06)
- PI of the National Security Agency Grant on *Censored Regression Quantile* (02/04-02/07)
- Statistician on the National Institute of Health Grant RO1-DC3684 *The Coordination of Respiration and Deglutition* (01/01-12/04)
- Co-Investigator, *Robustness in Generalised Semiparametric Mixed Models*, funded by the Research Grants Council of Hong Kong (08/03-07/06)
- PI of the National Security Agency Grant on *Regression Analysis* (02/02-02/04)
- PI of the National Science Foundation Award SBR-9617278 on *Constraints and Flexibility in Modeling* (09/97-08/01)

- PI of the National Security Agency Grant MDA904-96-1-0011 on *Robust methods* (02/96-02/98)
- Statistician, *Study on Seismic Rehabilitation Costs*, funded by Federal Emergency Management Agency through CERL (2002-2003)
- Statistician, *Child Welfare Studies* funded by the State of Illinois through the Children and Family Research Center, UIUC (2000-2001)
- Statistician, *Design and Analysis of Track Wear Tests* funded by the Association of American Railroads (2001-2002)
- University of Illinois Campus Research Board Awards on robust statistics (1994, 1996, 1997), computational statistics (1998), item response theory models (2000), and on bioinformatics (2002).

SELECTED CONSULTING EXPERIENCE:

- Consultant (2007-2008) to Millennium Pharmaceuticals, Inc. on design and analysis of clinical trials
- Consultant (2002-2003) to Archer-Daniels-Midland Company on microarray data analysis
- Consultant (2001-2003) to the Federal Commodity Futures Trading Commission (CFTC) on evaluating evidence of illegal trading
- Consultant (1995-2002) at the Argonne National Laboratory on residential energy consumption (sponsored by the U.S. Department of Energy)
- Consultant (1996) on the roofing data management for US Army Construction Engineering Research Laboratories
- Consultant (1995) for the Customer Potential Management Corporation at East Peoria, Illinois on modeling marketing data

SELECTED INVITED PRESENTATIONS SINCE 2001:

- Conference on Contemporary Methods of Data Analysis: Theory and Practice, Argentina, March 2001
- Department of Biostatistics, University of North Carolina at Chapel Hill, April 2001
- International Conference on Robust Statistics, Austria, July 2001
- Session Organizer for the 5th ICSA International Conference, Hong Kong, August 2001
- Department of Statistics, Columbia University, March 2002
- Invited speaker and Scientific Committee member, the 2nd International Conference on Robust Statistics, Vancouver, June 2002
- Department of Statistics, University of Virginia, December, 2002
- Invited speaker and Scientific Committee member, the 3nd International Conference on Robust Statistics, Antwerp, Belgium, July 2003
- BIRS Workshop on regularization in statistics, Banff International Research Station for Mathematical Innovation and Discovery, September, 2003
- Department of Statistics, North Carolina State University, March, 2004
- Organizing Committee member, the 6th ICSA Conference, Singapore, July 2004
- Joint Statistical Meetings, Toronto, August 2004
- Colloquium speaker "An Exciting Time for Statisticians" at the University of Virginia and Yale University, November, 2004

- Invited speaker at the 2004 Dysphagia Research Society on "Effects of Age, Gender, Bolus Volume and Viscosity on Acoustic Signals of Normal Swallowing", Montreal, October 2004
- Tutorial lecturer on "Quantile Regression" at the 2005 ENAR Meeting, March 2005
- Public School Lecture, "From Data to Discoveries" at NUS High School, Singapore, March 2, 2005
- Department of Mathematics, University of Maryland, March, 2005
- Department of Biostatistics, Columbia University, April, 2005
- Invited session organizer on robust statistics, the 25th European Meeting of Statisticians, Oslo, July, 2005
- Department of Biostatistics and Applied Mathematics, MD Anderson Medical Center, October, 2005
- School of Industrial and Systems Engineering, Georgia Tech, November, 2005
- The 5th Asian Conference on Statistical Computing, Hong Kong, December, 2005
- The TMS & AMS Joint International Conference, Taiwan, December, 2005
- Department of Statistics & Probability, Michigan State University, February, 2006
- Department of Information & Operations Management, University of Southern California, February, 2006
- Invited panelist at the Dysphagia Research Society Conference, Arizona, March, 2006
- Invited speaker at the Workshop on quantile regression, LMS method and robust statistics in the 21st century, Edinburgh, UK, June, 2006
- Invited speaker at the 2006 International Conference on Robust Statistics, Lisbon, Portgal, July, 2006
- Department of Statistics, University of Wisconsin at Madison, October 4, 2006
- Department of Mathematics, Washington University in St. Louis, October 19, 2006
- Department of Mathematics, Zhejiang University, China, September 13, 2006
- Invited lecture at the Millennium Pharmaceuticals, Inc., Boston, September 18, 2006
- Invited speaker at the Biostatistics Workshop, Institute for the Mathematical Science, Singapore, October 23, 2006.
- Department of Statistics, Wharton School, University of Pennsylvania, November 15, 2006
- Department of Statistics, Texas A&M University, December 15, 2006
- Department of Statistics, Yale University, February 5, 2007
- Department of Statistics, Oregon State University, February 26, 2007
- Invited speaker at the ENAR 2007 meeting, Atlanta, March 13, 2007
- Department of Statistics and Actuarial Science, University of Iowa, April 12, 2007
- Department of Statistics, Texas A&M University, May 17, 2007
- Invited Participant, Workshop on Statistical Analysis of Genetic Data, Banff, Canada, June 24-29, 2007
- Department of Biostatistics, Columbia University, July 12, 2007
- Mentor for the ASA Writers Workshop, Salt Lake City, July 29, 2007
- Keynote speaker at the International Conference on Frontiers of Statistics - High Dimensional Data, Kunming, China, August 13-15, 2007
- School of Public Health, Rochester University, September 6, 2007
- Department of Statistics, University of California at Los Angeles, October 2, 2007
- Department of Biostatistics, University of North Carolina at Chapel Hill, November 7, 2007
- Department of Mathematics and Statistics, Miami University, November 13, 2007

- Department of Operations Research and Financial Engineering, Princeton University, March 25, 2008
- Invited speaker at the first International Biopharmaceutical Statistics Conference, Shanghai, July 1, 2008
- Invited speaker at the International Workshop on Applied Probability (IWAP 2008), Compiegne, France, July, 2008
- Invited speaker at the Symposium on Advances in Statistics, in honor of the 65th birthday of Professor Zhidong Bai, Singapore, July 20, 2008
- Invited speaker and member of the Scientific Committee of the International Conference on Robust Statistics, Antalya, Turkey, September 8-12, 2008
- Keynote speaker at the Korean Statistical Society Fall Meeting, Seoul, October 31, 2008
- Department of Statistics, University of Michigan, November 7, 2008
- Invited speaker at the ENAR Meeting, San Antonio, March 17, 2009
- Distinguished lecturer at the 1st IMS Asia Pacific Rim Meeting, Seoul, July 1, 2009
- Colloquium speaker at the ASA Oregon Chapter Meeting, and Department of Mathematics, Portland State University, October 23, 2009
- Department of Statistics, University of Minnesota, November 19, 2009
- Invited speaker and discussant at Workshop on Statistical Frontiers, 2009 ISS, Taiwan, December, 2009
- Yale School of Public Health, Yale University, January 12, 2010
- Colloquium speaker, Department of Biostatistics, Harvard University, February 18, 2010
- Plenary speaker at the 2010 International Conference on Robust Statistics, July 1, 2010.

PUBLICATIONS

1. He. X, D.G. Simpson and S.L. Portnoy (1990): Breakdown Robustness of Tests, *Journal of the American Statistical Association*, vol. 85, no. 40, 446-452.
2. He, X. (1990): Estimating the root of a nonparametric regression function in a robust fashion, *The Australian Journal of Statistics*, vol. 32, no. 2, 217-226.
3. He, X., J. Jureckova, R. Koenker and S. Portnoy (1990): Tail Behavior of Regression Estimators and Their Breakdown Points, *Econometrica*, vol. 58, no. 5, 1195-1214.
4. He, X. (1990): Likelihood ratio test for discordancy with slippage alternatives, *Communications in Statistics, Theory and Methods*, vol. 19, no. 10, 3585-3594.
5. He, X. (1990): Common-sense monotonicity of robust estimators, *Metron*, vol. 48, 141-147.
6. He, X. (1991): A local breakdown property of robust tests in linear regression, *Journal of Multivariate Analysis*, vol. 38, no. 2, 294-305.
7. He, X. and D.G. Simpson (1992): Robust Direction Estimation, *The Annals of Statistics*, vol. 20, no. 1, 351-369.
8. He, X. (1992): Robust Statistics of Directional Data: A Survey. *Nonparametric Statistics and Related Topics*, edited by A.K. Md.E. Saleh, North Holland, New York, 87-96.
9. He, X. and S.L. Portnoy (1992): Reweighted LS estimators converge at the same rate as the initial estimator. *The Annals of Statistics*, vol. 20, no. 4, 2161-2167.
10. He, X. and D.G. Simpson (1993): Lower Bounds for Contamination Bias: Globally Minimax Versus Locally Linear Estimation. *The Annals of Statistics*, vol. 21, 314-337.
11. Chan, Y.M. and He, X. (1994): On Median-type Estimators of Direction for the von Mises- Fisher Distributions. *Biometrika*, vol. 80, 869-875.

12. Chan, Y.M. and He, X. (1994): A simple and competitive estimator of location. *Statistics and Probability Letters*, vol. 19, 137-142.
13. He, X. and P. Shi (1994): Convergence rate of B-spline estimators of nonparametric conditional quantile functions. *J. Nonparametric Statistics*, vol. 3, 299-308.
14. Markatou, M. and He, X. (1994): Bounded influence, high breakdown point testing procedures in linear models. *J. Amer. Statist. Assoc.*, vol. 89, 543-549.
15. He, X. (1994): Breakdown versus efficiency, your perspective matters, as Guest Editor for *Statistics and Probability Letters*, vol. 19, no. 5, 357-360.
16. He, X. and G. Wang (1995): Law of the Iterated Logarithm and Invariance Principle for M- estimators. *Proceedings of the American Mathematical Society*, vol. 123, no. 2, 563-573.
17. Cui, H. and He, X. (1995) Consistency of semiparametric estimation of elliptic densities, *Acta Mathematica Sinica, new Series*, vol. 11, 44-58.
18. He, X. and G. Wang (1996): Cross-checking using the minimum volume ellipsoid estimator. *Statistica Sinica*, vol. 6, no. 2, 367-374.
19. He, X. and P. Shi (1996): Asymptotics for M-type regression splines with auxiliary scale estimation, *Sankhya-A*, 452-461.
20. He, X. and Q.M. Shao (1996): Bahadur efficiency and robustness of Studentized score tests, *Annals of Instit. Statist. Math*., vol. 48, no. 2, 295-314.
21. He, X. and P. Shi (1996): Bivariate tensor-product B-splines in a partly linear model, *Journal of Multivariate Analysis*, vol. 58, no. 2, 162-181.
22. He, X. and Q.M Shao (1996): A general Bahadur representation of M-estimators and its application to linear regression with nonstochastic designs, *Annals of Statistics*, vol. 24, no. 6, 2608-2630.
23. He, X. and G. Wang (1997): A qualitative robustness of S* estimators of multivariate location and dispersion, *Statistica Neerlandica*, vol. 51, no. 3, 257-268.
24. He, X. and L. Shen (1997): Linear regression after spline transformation, *Biometrika,* vol. 84, no. 2, 474-481.
25. He, X. (1997): Quantile curves without crossing, *The American Statistician*, vol. 51, no. 2 186- 192.
26. He, X. and G. Wang (1997): Convergence of depth contours for multivariate datasets, *Annals of Statistics*, vol. 25, no. 2, 495-504.
27. He, X., P. Ng and S. Portnoy (1998): Bivariate quantile smoothing splines, *Journal of the Royal Statistical Society, Series B*, vol. 60, Part 3, 537-550.
28. He, X. and S. Portnoy (1998): Asymptotics of the deepest line, chapter 5 of *Applied Statistical Science III*, pp. 71-81, Nova Science Publishers, Inc.
29. Chen, L. and He, X. (1998): Test of normality for censored data, chapter 8 of *Applied Statistical Science III*, pp. 113-125, Nova Science Publishers, Inc.
30. He, X. and P. Shi (1998): Monotone B-spline smoothing, *Journal of the American Statistical Association*, Vol. 93, No. 442, 643-650.
31. Simpson, D.G., He, X. and Y.T. Liu (1998): Comments on "M-type smoothers with edge preserving properties", *Journal of the American Statistical Association*, Vol. 93, No. 442, 544- 547.
32. He, X. and P. Ng (1999): Quantile Splines with several covariates, *Journal of Statistical Planning and Inference*, vol. 75(2), 343-352.

33. He, X. and P. Ng: (1999): COBS: Constrained Smoothing via Linear Programming, *Computational Statistics*, vol. 14, 315-337.
34. He, X. (1999): Comments on "regression depth", *Journal of the American Statistical Association*, vol. 94, 403-404.
35. He, X. and W.K. Fung (1999): Method of Medians for life time data with Weibull models, *Statistics in Medicine*, vol. 18, 1993-2009.
36. Bai, Z.D. and He, X. (1999): Asymptotic distributions of the maximal depth estimators for regression and multivariate location. *Annals of Statistics*, vol. 27, 1616-1637.
37. Galfalvy, H.C., He, X., and D. G. Simpson (1999): Conditional quantile curves with an application to infrastructure studies. *American Statistical Association 1999 Proceedings of the Section on Statistical Consulting*, 37-41.
38. He, X. and H. Liang (2000): Quantile regression estimates for a class of linear and partially linear errors-in-variables models, *Statistica Sinica*, vol. 10, 129-140.
39. He, X. and W.K. Fung (2000): High breakdown estimation for multiple populations with applications to discriminant analysis, *Journal of Multivariate Analysis*, vol. 72, 151-162.
40. He, X. and Q.M. Shao (2000): On parameters of increasing dimension, *Journal of Multivariate Analysis*, vol. 73, 120-135.
41. He, X. (2000): Comment on "Multivariate L-estimation", *Test*, vol. 8, 297-301.
42. He, X., D.G. Simpson, and G.Y. Wang (2000): Breakdown points of t-type regression estimators. *Biometrika*, Vol. 87, 675-687.
43. He, X. and S. Portnoy (2000): Some asymptotic results on bivariate quantile splines, *Journal of Statistical Planning and Inference*, Vol. 91, No. 2, 341-349.
44. He, X. and W.K. Fung (2000): Discussion of "Dimension Reduction and Visualization in Discriminant Analysis", *Australian & New Zealand Journal of Statistics*, Vol. 43, no. 2, 190- 193.
45. Portnoy, S. and He, X. (2000): A robust journey in the new millennium, *Journal of the American Statistical Association*, Vol. 95, No. 452, 1331-1335.
46. He, X., Z.W. Shen and L.X. Shen (2001): A data-adaptive knot selection scheme for fitting splines. *IEEE Signal Processing Letters*, Vol. 8, No. 5, 137-139.
47. Monrad, D., W. F. Stout, B. A. Bailey, J.R. Fryxell, R.L. Gould, He, X., V.R. Plessner, and L. Roussos (2001): Statistics, the Craft of Data Collection, Description, and Inference (Second Edition). Textbook from Mobius Communications.
48. He, X. and M. Kim (2002): On Marginal Estimation in a Semiparametric Model for Longitudinal Data with Time-independent Covariates. *Metrika* (special issue), Vol. 55, 67-74.
49. Huang, B., Bachmann, K., He, X., Chen, R., Mcallister, J. and Wang, T. (2002): Inappropriate prescriptions for the aging population of the United States: an analysis of the National Ambulatory Medical Care Survey, 1997, *Pharmacoepidemiology and Drug Safety*, Vol. 11, 127-134.
50. He, X., Zhu, Z.Y. and Fung, W.K. (2002): Estimation in a Semiparametric Model for Longitudinal Data with Unspecified Dependence Structure, *Biometrika*, Vol. 89, No. 3, 579- 590.
51. Fung, W.K., He, X., Liu, L. and Shi, P.D. (2002): Dimension reduction based on canonical correlation, *Statistica Sinica*, Vol. 12, 1093-1114.

52. Cui, H., He, X. and Zhu, L. (2002): Asymptotics of De-noised Least Squares Estimators, *Statistica Sinica*, Vol. 12, 1191-1206.
53. He, X. and Hu, F. (2002): Markov chain marginal bootstrap, *J. American Statistical Association*, Vol. 97, no. 459, 783-795.
54. Fung, W.K., Zhu, Z.Y., Wei, B.C. and He, X. (2002): Influence Diagnostics and Outlier Tests for Semiparametric Mixed Models, *J. Royal Statistical Society, Series B*, Vol. 64(3), 565-579.
55. He, X., Ng, W.K. and Shi, J. (2003): Marginal versus Joint Box-Cox Transformation with Applications to Percentile Curve Construction for IgG Subclasses and Blood Pressures, *Statistics in Medicine*, Vol. 22, 397-408.
56. He, X. (2003): Robust Tests in Statistics – a Review of the Past Decade. For a special issue of *Estatistica*, Vol. 54, 29-46.
57. Chen, M.H., He, X., Shao, Q.M., and Xu, H. (2003): A Monte Carlo Gap Test in Computing HPD Regions, in *Development of Modern Statistics and Related Topics - In celebration of Professor Yiao-Ting Zhang's 70th Birthday,* Series in Biostatistics, Vol. 1, World Scientific, Singapore.
58. Zhu, Z.Y., He, X. and Fung, W.K. (2003): Local influence analysis for penalized Gaussian likelihood estimators in partially linear models. *Scandinavian Journal of Statistics*, Vol. 30, 767-780.
59. Bai, Z.D. and He, X. (2003): A Chi-square test for dimensionality with non-Gaussian Data. *Journal of Multivariate Analysis*, Vol. 88, 109-117.
60. He, X., Fu, B. and Fung, K.W. (2003): Median Regression of Longitudinal Data. *Statistics in Medicine*, Vol. 22, 3655-3669.
61. Cui, H., He, X. and Ng, K.W. (2003): Asymptotic Distributions of Principal Components Based on Robust Dispersions. *Biometrika*, Vol. 90, 954-966.
62. He, X. and Zhu, L. (2003): A Lack-of-fit Test for Quantile Regression, *Journal of the American Statistical Association*, Vol. 98, 1013-1022.
63. He, X., Cui, H. and Simpson, D.G. (2004): Longitudinal Data Analysis Using t-type Regression, *Journal of Statistical Planning and Inference*, Vol. 122, 253-269.
64. Zuo, Y., H. Cui and He, X. (2004): On the Stahel-Donoho Estimator and Depth-weighted Means of Multivariate Data, *Annals of Statistics*, Vol. 32, No. 1, 167-188.
65. Cui, H., He, X. and Ng, K.W. (2004): M-estimation for linear models with spatially-correlated errors, *Statistics and Probability Letters*, Vol. 66, No. 4, 383-393.
66. Lin, N., Bailey, B., He, X. and Buttlar, W. (2004): Adjustment of Measuring Devices with Linear Models, *Technometrics*, Vol. 46, No. 2, 127-134.
67. He, X. and Portnoy, S. (2004): Comment on Location-Scale Depth, Jo*urnal of the American Statistical Association*, Vol. 99, No. 468, 973-976.
68. He, X. and Ng, K.W. (2004): Percentile charts for correlated measures. In *Encyclopedias of Biopharmaceutical Statistics*, edited by Shein-Chung Chow, Dekker.com.
69. Kocherginsky, M., He, X. and Mu, Y. (2005): Practical Confidence Intervals for Regression Quantiles, *Journal of Computational and Graphical Statistics*, Vol. 14, no. 1, 41-55.
70. Perlman, A.L., He, X., Barkmeier, J., and Van Leer, E. (2005): Bolus Location Associated with Videofluoroscopic and Respirodeglutometric Events. *Journal of Speech, Language, and Hearing Research*, Vol. 48, no. 1, 21-33.

71. Huang, B., Martin, S.J., Bachmann, K.A., He, X., Reese, J.H., Wei, Y., Iwuagwu C., Mao, Z., and Xin, H. (2005). A nationwide survey of physician office visits found that inappropriate antibiotic prescriptions were issued for bacterial respiratory tract infections in ambulatory patients. *Journal of Clinical Epidemiology*, Vol. 58, 414-420.
72. He, X. (2005): Discussion on "Breakdown and Groups", *Annals of Statistics*, Vol. 33, 998-1000.
73. He, X., Fung, W.K., and Zhu, Z.Y. (2005): Robust Estimation in Generalized Partial Linear Models for Clustered Data. *Journal of the American Statistical Association*, Vol. 100, 1176- 1184.
74. Wang, H.X., He, X., Band, M., Wilson, C. and Liu, L. (2005): A study of inter-lab and inter-platform agreement of DNA microarray data. *BMC Genomics*, Vol. 6, Art. No. 71.
75. Boente, G., He, X. and Zhou, J. (2006): Robust Estimates in Generalized Partially Linear Models. *Annals of Statistics*, Vol. 34, 2856-2878.
76. Wei, Y., Pere, A., Koenker, R. and He, X. (2006): Quantile Regression Methods for Reference Growth Charts. *Statistics in Medicine*, Vol. 25, no. 8, 1369-1382.
77. Wei, Y. and He, X. (2006): Conditional Growth Charts (with discussions). *Annals of Statistics*, Vol. 25, no. 8, 1369-1382.
78. Lin, N. and He, X. (2006): Robust and efficient estimation under data grouping. *Biometrika*, Vol. 93, no. 1, 99-112.
79. Zuo, Y. and He, X. (2006): On the limiting distributions of multivariate depth-based rank sum statistics and related tests. *Annals of Statistics*, Vol. 34, 2879-2896.
80. Lin, G., He, X., Hanlee, J., Shi, L., Davis, R.W., and Zhong, S. (2006): Reproducibility Probability Score: Incorporating Measurement Variability across Laboratories for Gene Selection. *Nature Biotechnology*, Vol. 24(12), 1476-1477 (with online supplementary material)
81. Hu, J and He, X. (2007): Enhanced quantile normalization of microarray data to reduce loss of information in the gene expression profile. *Biometrics*, Vol. 63, No. 1, 50-59.
82. Mu, Y. and He, X. (2007): Power Transformation towards a Linear Regression Quantile. *Journal of American Statistician Association*, Vol. 102, 269-279.
83. Wang, H.X. and He, X. (2007): Detecting Differential Expressions in GeneChip Microarray Studies: A Quantile Approach. *Journal of the American Statistical Association*, Vol. 102, 104- 112.
84. Kocherginsky, M. and He, X. (2007). Extensions of the Markov Chain Marginal Bootstrap. *Statistics and Probability Letters*, Vol. 77, No. 12, 1258-1268.
85. Cui, H., He, X. and Liu. L. (2007). Testing for Additivity with B-Splines. *Science in China Series A: Mathematics*, Vol. 50, No. 6, 841-858.
86. He, X. and Hu, F. (2007). Comments on 'Implementation of Estimating-Function Based Inference Procedures with MCMC Samplers' by Tian, Liu and Wei. *Journal of the American Statistical Association*, Vol. 102, no. 479, 889-890.
87. Hu, J., He, X., Baggerly, K.A., Coombes, K. R., Hennessy, B., and Mills, G. B. (2007). Nonparametric Quantification of Protein Lysate Arrays. *Bioinformatics*, Vol. 23, No. 15, 1986- 1994.
88. Zhou, T. and He, X. (2008). Three-step Estimation in Linear Mixed Models with Skew-t Distributions. *Journal of Statistical Planning and Inference*, Vol. 138 (6), 1542-1555.
89. Henson, R., Roussos, L., Douglas, J. and He, X. (2008): Cognitive diagnostic attribute level discrimination indices. *Applied Psychological Measurement*, Vol. 32, No. 4, 275-288.

90. Wang, H. and He, X. (2008). An Enhanced Quantile Approach for Assessing Differential Gene Expressions. *Biometrics*, Vol. 64(2), 449-457.
91. Chen, C., He, X., and Wei, Y. (2008). Lower Rank Approximation of Matrices Based on Fast and Robust Alternating Regression. *Journal of Computational and Graphical Statistics*, Vol. 17, no. 1, 186-200.
92. Zhou, J. and He, X. (2008). Dimension Reduction Based on Constrained Canonical Correlation and Variable Filtering. *Annals of Statistics*, Vol. 36, no. 4, 1649-1668.
93. Noe, D. and He, X. (2008). Partially Bayesian Variable Selection in Classification Trees. *Statistics and its Interface*, Vol. 1, 155-167.
94. Thorpe MP, Jacobsen EH, Layman DK, He X, Kris-Etherton PM, Evans EM (2008). A diet high in protein, dairy and calcium attenuates bone loss over 4 months weight loss and 8 mo weight maintenance relative to a conventional high carbohydrate diet in adults. *Journal of Nutrition*, Vol. 138, 1096-1100.
95. Zhu, Z.Y, Fung, W.K., and He, X. (2008). A note on the local asymptotics of marginal regression splines with longitudinal data. *Biometrika*, Vol. 94, no. 4, 907-917.
96. Hu, F., Zhang, L.X., and He, X. (2009). Efficient Randomized Adaptive Designs. *Annals of Statistics*, Vol. 37, No. 5A, 2543-2560.
97. Broglio, S. P., Sosnoff, J., Shin, S., He, X., Alcaraz, C., and Zimmerman, J. (2009). Helmet impacts during high school football: A biomechanical assessment. *Journal of Athletic Training*, Vol. 44(4), 342-349.
98. Hu, Y., He, X., Tiao, J., and N. Z. Shi (2009). Modeling and prediction of children's growth data via functional principal component analysis. *Science in China, Series A*, Vol. 52, No. 6, 1342-1350.
99. Yang, S.K., Long, C., Miller, A., He, X., Yang, Y., Wuebbles, D., and Tiao, G. (2009). The modulation of natural variability on the trend analysis of the updated cohesive SBUV(2) total ozone. *International Journal of Remote Sensing*, Vol. 30, 3975-3986.
100. Feng, X. and He, X. (2009). Inference on low-rank data matrices with applications to microarray data. *Annals of Applied Statistics*, Vol. 3, No. 4, 1634-1654.
101. Shi, N.Z., He, X., and Tao, J. (2009). Understanding Statistics and Statistics Education: A Chinese Perspective. *Journal of Statistics Education*, Volume 17, Number 3, www.amstat.org/ publications/jse/v17n3/shi.html.
102. Hu, J, He, X., Cote, G.J., and Krafe, R. (2009). Singular Value Decomposition-based Alternative Splicing Detection. *Journal of the American Statistical Association*, Vol. 104, No. 487, 944-953.
103. Hong, F., Fang, F, He, X., Cao, X., Chipperfield, H., Xie, D., Wong, W. H., Ng, H. H., and Zhong, S. (2009). Dissecting early differentially expressed genes in a mixture of differentiating embryonic stem cells. *PLoS Computational Biology*, Vol. 5(12), doi:10.1371/journal.pcbi.1000607.
104. Hong, H. G., and He, X. (2010). Prediction of Functional Status for the Elderly Based on a New Ordinal Regression Model. *Journal of the American Statistical Association*, Vol. 105, No. 491, 930-941.
105. He, X., Xue, H., and Shi, N.Z. (2010). Sieve Maximum Likelihood Estimation for Doubly Semiparametric Zero-Inflated Poisson Models. *Journal of Multivariate Analysis*, Volume 101, Issue 9, 2026-2038.

106. He, X. and Zhou, J. (2010). Discussion of "Envelope models for parsimonious and efficient multivariate linear regression", *Statistica Sinica*, Vol. 20, No. 3, 971-978.
107. Broglio, S. P., Schnebel, B., Sosnoff, J.J., Shin, S., Feng, X., He, X., Zimmerman, J. (2010). The biomechanical properties of concussions in high school football. *Medicine & Science in Sports & Exercise*, Vol. 42, 2064-1071.
108. He, X., Hsu, Y. H., and Hu, M. (2010). Detection of Treatment Effects by Covariate-adjusted Expected Shortfall. *Annals of Applied Statistics*, Vol. 4, No. 4, 2114-2125.
109. Li, C., Wei, Y., Chappell, R. J. and X. He (2011). Bent Line Quantile Regression with application to an allometric study of land mammals' speed and mass. *Biometrics*, Vol. 67, 242- 249.
110. Wang, X., Gui, J., and He, X. (2011). Finding the minimal set for collapsible graphical models. *Proceedings of the American Mathematical Society,* Vol.139, 361-373.
111. Guo, J., Xu, P.F., and He, X. (2011). An Improved Iterative Proportional Scaling Procedure for Gaussian Graphical Models. *Journal of Computational and Graphical Statistics*, Vol. 20, 417-431.

SOFTWARE:

- He, X. and Ng, P. (2001): Constrained B-spline Smoothing (COBS), Fortran functions with S-Plus interface, and in CRAN for R users.
- Kocherginsky, M. and He, X. (2003): MCMB Confidence Intervals for Regression Quantiles, currently available as a contributed package *rqmcmb2* in CRAN for R users. The same method is included in SAS PROC QUANTREG under the *mcmb* option.

Christian E. Lindhjem

EDUCATION

1987	PhD, Chemical Engineering, Rensselaer Polytechnic Institute
1981	MS, Chemical Engineering, University of Michigan
1980	BS, Chemical Engineering, Rose-Hulman Institute of Technology
1980	BS, Chemistry, Rose-Hulman Institute of Technology

EXPERIENCE

Dr. Christian E. Lindhjem, a Senior Consultant at ENVIRON, is an expert on emissions from highway and nonroad vehicles and engines and fuels used in those engines. With ENVIRON for 12 years and with EPA's Office of Transportation and Air Quality (previously named Office of Mobile Sources) for 8 years, Dr. Lindhjem has worked on on-road and off-road mobile source regulation development, emission measurements and analysis, emission control strategies, and emission inventory modeling including regulated pollutants and chemical compositional analysis to estimate toxic emissions and other components of concern. Dr. Lindhjem has evaluated and continues to consult on a broad range local and national mobile source emission issues including both emission estimates and potential emissions reductions from a variety of planned or demonstrated control strategies for on-road vehicles and off-road equipment (such as construction equipment, locomotive, marine vessels, and other harbor sources) and heavy-duty diesel vehicles. He works with local, regional and national officials and private clients to improve on-road and off-road (including agricultural, commercial marine, locomotive, construction and mining, recreational marine, and lawn and garden equipment types) emission inventories for regional evaluation as well as individual facilities including ports and rail yards. His other national experience included assistance for EPA in the development of the NONROAD model and a national commercial marine emission inventory and spatial allocation for deep draft vessels. Dr. Lindhjem holds a Ph.D. degree in Chemical Engineering from the Rensselaer Polytechnic Institute and a M.S. degree in Chemical Engineering from the University of Michigan.

- Commercial Marine, Rail, and Intermodal Emissions Evaluations: Conducted in depth analysis and method formulation to revise locomotive and commercial marine emissions for ports and railroads, states, and national agencies including those for the States of Arkansas, Illinois, Indiana, Michigan, Ohio, Wisconsin, and Wyoming and the metropolitan areas of Kansas City and Boise. Prepared a grid scale national emissions inventory from Category 3 (ocean-going vessels) commercial marine engines for EPA combining near port and open-ocean emissions. Other projects include emissions and control strategy evaluations for the Ports of Long Beach, San Francisco, Oakland, San Diego and others, including the emission reduction potential from shore power and other control strategies. Evaluated averaging strategies to comply with the Canadian and U.S. Emission Control Area (ECA) designation for ocean-going vessels using highly resolved activity and emissions calculations. Conducted detailed rail yard activity and emissions evaluations in California and Kansas including locomotive and other off-road and on-road source categories within the yards.
- On-Road Emission Inventories: Managed projects to estimate metropolitan area and project level emissions estimates using the EPA MOVES2010 and California ARB EMFAC models. Managing projects to incorporate CONCEPT link level modeling for Detroit, Atlanta, and Philadelphia including MOVES2010 model adjustments and vehicle volume and mix of heavy and light-duty vehicles adjusted by time of data and road type. Performed evaluations of the on-road vehicle mix and vehicle weight data from Automatic Traffic Recorders to better estimate on-road emissions for EPA, Illinois, Ohio, Wisconsin, and Minnesota. Led programs to develop new strategies and uses for

alternative data to evaluate on-road vehicle emissions for EPA and the Coordinating Research Council in preparation of the MOVES model.

- On and Off-Road Mobile Source Control Strategies: Provides detailed technical analyses and planning and day to day support of on-road and off-road mobile source control strategies for local and state government agencies including the Houston-Galveston, Dallas-Ft. Worth, and East Texas nonattainment areas' SIPs, for the Houston-Galveston Area Council (HGAC), the East Texas Council of Governments (ETCOG), North Central Council of Governments (NCTCOG), and the Texas Commission on Environmental Quality (TCEQ). Also assisting in revisions/updates to the mobile source emission inventory in Texas and Las Vegas. Providing the plan and technical guidance for the HGAC Voluntary Mobile Source Emission Reduction Program (VMEP) including marketing and online database and interactive programs, and reviews and evaluation of the Texas Emission Reduction PLAN (TERP) for the Houston Advanced Research Center (HARC).
- Off-Road Emission Inventories: Provided service to state agencies by improving air emissions inventories from all nonroad mobile sources. Conducted in-depth locomotive and other emissions evaluations for BNSF railyards in California in preparation for dispersion modeling impact analysis on the local community. For the Lake Michigan Air Directors Consortium (LADCO), a plan was developed and implemented for revising and developing activity estimates for all off-road emission sources, incorporate such estimates in the emissions inventory, and spatially allocate those emissions. Innovative methods were used for commercial marine to quantify the Lake and river commercial marine emissions, locomotive emissions from large and small railroads, and a unique method for recreational marine spatial allocation specifically. Prepared revised emissions inventories some categories or all nonroad (including aircraft, commercial marine, and locomotive sources) for the States of Arizona, Texas, Arkansas, Oklahoma, Wyoming, Wisconsin, Illinois, Indiana, Michigan, Ohio, New Jersey, and the Western Regional Air Partnership (WRAP). Assisted EPA in improvements to the NONROAD model for off-road emissions estimation including for instance detailed technical comparisons of EPA and CARB models for estimating emissions from nonroad mobile sources and providing revisions to modeling structure, activity data inputs and assumptions, and emission factors for nonroad spark-ignition and compression ignition engines. Another example of activity includes providing technical guidance to Northeast States for Coordinated Air Use Management (NESCAUM) for the development of surveys and analysis of survey data to derive improved estimates of construction equipment activity in the Northeast U.S.
- Scientific Emissions Research and Analysis: Performed scientific evaluation of emission modeling methods such as critically reviewing the project plan and available data for EPA's new MOVES model in an extensive report (E-68) for the Coordinating Research Council (CRC). This report outlined the general method, specifics of data handling, and many other areas of interest in implementing this new modeling approach for on-road mobile sources. The CRC project followed the previous work for EPA, which developed and executed a method using on-board emissions monitor data to estimate emissions from onroad and offroad vehicles and equipment for the Office of Transportation and Air Quality (OTAQ) of EPA. Driving behavior and emissions were related to the physical and operating parameters experienced of the vehicle or engine. Scientific evaluation of the effects of humidity and temperature on mobile source emissions was conducted and incorporated into Texas emissions for HARC and TCEQ. A scientific assessment of the emissions response from hybrid-electric urban buses for the New York State Energy Research and Development Agency (NYSERDA).
- Chemical Constituents of Mobile Source Emissions: Performed a literature review and evaluation of the chemical constituents including potential toxic components of mobile source emissions for the EPA. Estimated the emission reduction potential and impacts on air quality and toxicity from replacing diesel with biodiesel fuel in heavy-duty diesel vehicles for the Department of Energy's National Renewable Energy Laboratory.

- Mobile Source Emissions Air Quality Impact Evaluation: Providing limited and innovative plans to evaluate and mitigate mobile source emissions at a project level. For instance, the City of Hawthorne, California, provided an evaluation of the impact and provided an innovated contracting method to mitigate air quality effects for garbage collection vehicles. Provided technical support for estimating on-road emissions and air quality impacts for the expansion of an amphitheater and casino including preparing a report for the Environmental Impact Statement and an assessment of General Conformity.

Prior to joining ENVIRON, Dr. Lindhjem held the following positions:

Engineer, U.S. Environmental Protection Agency (EPA), National Vehicle and Fuel Emissions Laboratory, Office of Mobile Sources, Ann Arbor, MI (1990 - 1998), 8 years

Nonroad Mobile Emissions

- Evaluation of rail and port intermodal yard emissions including emission activity from ships, locomotive, cargo handling equipment, and on-site truck movements.
- Responsible for the theoretical framework behind EPA's NONROAD model, the next inventory modeling tool for nonroad engines, analogous to EPA's MOBILE model for highway vehicles, including nonroad industrial, construction, commercial, residential, and marine engines and vehicles.
- Developed the input estimates for the EPA's NONROAD model for population, activity, load factor, average life, and compression and spark-ignition emission factors for nonroad engines as well as estimates for the hydrocarbon speciation.
- Project engineer on a variety of test programs measuring emissions from nonroad engines, specifically construction and agricultural diesel engines, outboard and inboard recreational marine motors, and lawn & garden engines.
- Evaluated of the effect of test cycle on various diesel and gasoline nonroad engines through regulatory and nonregulatory steady-state and transient comparisons.
  Developed a methodology to estimate emissions from commercial marine engines.

Highway Mobile Emissions Model

- Incorporated MOVES into emission inventory development tools such as SMOKE, CONCEPT, and project level emissions development.
- Have provided emission estimates for light-duty and heavy-duty vehicles exhaust and evaporative emissions, and fuel effects for MOBILE4.1, MOBILE5, and MOBILE6.
- Specific areas of emissions modeling were estimating the effects of reformulated gasoline, evaporative emissions, and diesel engines emissions for heavy-duty highway trucks and buses.
- Other modeling efforts include analyzing chemical characterizations of exhaust and evaporative emissions to provide estimates currently used in EPA's Speciate Database.
- Managed up to $1.4 million in contract funding per year.
- Research included lean NOx-reduction catalyst evaluations, test cycle comparisons for highway and nonroad engines, and novel emission modeling techniques for diesel engines.

Engine Testing Project Management

- Managed the heavy-duty, marine, evaporative, and light-duty chassis testing programs with five technicians, two engineers, and a budget of about $500k per year.
- Multidisciplinary approach combining mechanical, electrical, computer, and chemical analysis.
- Supplied test results for a variety of EPA projects using a variety of engines and vehicles; most notably, nonroad and highway diesel, inboard and outboard marine, and lawn and garden engines.

Clean Fuels Development

- Developed the Reformulated Gasoline Simple Model; evaluated the effect of fuel oxygen and volatility control on primary and toxic pollutants from gasoline vehicles.
- Team member for the Reformulated Gasoline Complex Model and Final Rulemaking providing estimates for effects of fuel parameter changes on primary and toxic pollutants.
- Conducted testing programs investigating the effects of reformulated fuels on vehicles and nonroad engines.
- Facilitated fuel additive emission testing and registration.

Research Engineer, Westvaco Corporation, (1988 - 1990), Laurel Research Center, Laurel, MD

- Designed coating formulations and products for the fine papers division.
- Successfully quantified subjective criteria of product quality and determined the mechanisms of in-use deterioration of our products during printing.
- Determined rheological failures in the production of coated papers and developed novel coating formulations for improved production.
- Managed 2 technicians and interactions with research and production.

EPA AWARDS

- EPA Bronze Medal for Commendable Service for Nonroad Engine Emission Controls Development, 1995
- EPA Bronze Medal for Commendable Service for Highway Heavy-Duty Engine Emission Reduction, 1995
- EPA Science Achievement Award in Air Quality, 1992
- EPA Silver Medal for Superior Service for Clean Fuels Development, 1991

TEACHING EXPERIENCE

University of Michigan, (1991), “Pollution Control for Chemical Engineers,” Department of Chemical Engineering
Rensselaer Polytechnic Institute, (1985-1987), “Air Pollution Control,” Department of Chemical Engineering

PROFESSIONAL MEMBERSHIPS

Air and Waste Management Association
National Cooperative Highway Research Program Panelist
EPA’s Mobile Source Technical Review Subcommittee for Nonroad Engine Emissions

PUBLICATIONS AND PRESENTATIONS

Lindhjem C.E. 2010 "Use of MOVES2010 in Link Level On-Road Vehicle Emissions Modeling Using CONCEPT-MV," C. E. Lindhjem, A. DenBleyker, M. Jimenez, J. Haasbeek, A. K. Pollack, ENVIRON International Corporation, CA; Z. Li, Clark County, Las Vegas, NV. 19th International Emission Inventory Conference, San Antonio, Texas, September 27 - 30, 2010.

Lindhjem C.E. 2010 "Development of Drivers and Post-Processing Scripts to Incorporate MOVES2010 Emission Factors with the Smoke Emissions Model,"C. E. Lindhjem, A. DenBleyker, M. Jimenez and A. K. Pollack, ENVIRON International Corporation. 19th International Emission Inventory Conference, San Antonio, Texas, September 27 - 30, 2010.

Lindhjem, C. 2009. “Mobile Source Particulate And Semi-Volatile Organic Carbon Ambient Modeling.” Presented at the 18th International Emissions Inventory Conference, Baltimore, MD, April 15, 2009.

Lindhjem, C. 2009 “Mobile Source Emissions: Adjustments to MOBILE6,” 19th CRC On-Road Vehicle Emissions Workshop, March 23-25, 2009.

Lindhjem, C. 2008. “Intermodal Yard Activity and Emissions Evaluations.” Presented at the 17th International Emissions Inventory Conference, Portland, OR. June.

Lindhjem, C.E. and Russell, J., 2006. “Development Of Gridded Ocean-Going Vessel Emission Inventories,” Presented at the Air and Waste Management Association Emission Inventory Conference, New Orleans, LA, May 17.

Lindhjem, C.E. and Shepard S. 2005. “Estimation and Effects of Vehicle Mix on On-Road Emission Estimates,” Air and Waste Management Association Emission Inventory Conference, Las Vegas, Nevada, April 14.

Lindhjem, C.E. and Chan L-M. 2004. “Emission Control Technologies and Programs for Heavy-Duty Diesel Vehicle Fleets in North America,” Paper No. 371. Air and Waste Management Association Annual Meeting, Indianapolis IN.

Lindhjem, C.E., Chan, L-M., Pollack, A.K., and Kite C. 2004. “Applying Humidity and Temperature Corrections to On and Off-Road Mobile Source Emissions,” Air & Waste Management Association Emission Inventory Conference, St. Petersburg Florida.

Lindhjem, C.E., T. Stockenius. 2002. “On-Board Emissions Data Analysis,” Presented at the12th CRC On-Road Vehicle Emissions Workshop, San Diego, California, April.

Lindhjem, C.E., A.K. Pollack, R. Chi. 2001. “Comparison Of Highway Mobile Source Emissions Inventory From MOBILE1 through MOBILE6,” 11th CRC On-Road Vehicle Emissions Workshop, San Diego, California, March.

Lindhjem, C.E., and D.A. Guerrieri. 1998. “Evaluation of Lean NOx Reduction Catalysts for Controlling Emissions from Diesel Engines”, *Environmental Progress*, Spring, 1998, page 48.

Lindhjem, C.E., D.M. Swain, C.J. Jackson, and G.J. Hoffman. 1998. “A Method for Comparing Transient NOx Emissions with Weighted Steady State Test Results,” *Society of Automotive Engineers, SAE-980408.*

Lindhjem, C.E., D.J. Korotney, V. Rao, and M.S. Sklar. 1995. “Reformulated Gasoline Effects on Exhaust Emissions: Phase III: Investigation on the Effects of the Oxygenate ETBE, Sulfur, Olefins, Volatility, and Aromatics and the Interactions Between Olefins and Volatility or Sulfur,” *Society of Automotive Engineers, SAE-950782.*

Lindhjem, C.E. 1995. “The Effect of Gasoline Reformulation and Sulfur Reduction on Exhaust Emissions from Post-1983 and Pre-1990 Vehicles,” *Society of Automotive Engineers, SAE-950778.*

Lindhjem, C.E., S.C. Mayotte, V. Rao, and M.S. Sklar. 1994. “Reformulated Gasoline Effects on Exhaust Emissions: Phase II: Continued Investigation of the Effects of Fuel Oxygenate Content, Volatility, Sulfur, Olefins, and Distillation Parameters,” *Society of Automotive Engineers, SAE-941974.*

Lindhjem, C.E., S.C. Mayotte, V. Rao, and M.S. Sklar. 1994. “Reformulated Gasoline Effects on Exhaust Emissions: Phase I: Initial Investigation of Oxygenate, Volatility, Distillation and Sulfur Effects,” *society of Automotive Engineers, SAE-941973.*

Lindhjem, C.E., D.J. Korotney. 1993. “Running Loss Emissions from Gasoline-Fueled Motor Vehicles,” *Society of Automotive Engineers, SAE-931991.*

Lindhjem, C.E. 1991. “The Particle Packing and Shape Effects on the Rheological Characteristics of Paper Coating Pigments,” *Proceedings of the 1991 TAPPI Coating Conference*, Montreal, Quebec, Canada, TAPPI Press, Atlanta, Georgia.

Lindhjem, C.E., E.R. Altwicker. 1988. “Absorption of Gases by Drops,” *AICHE Journal*, v. 34. pp. 329-332.

SELECTED REPORTS

Lindhjem, C.E. and Sturtz, T.M., “Development of Emission Estimates for Locomotives in the Kansas City Metropolitan Statistical Area,” Prepared for Kansas Department of Health and Environment (KDHE), May 13, 2010.

Parker, L., Lindhjem, C.E. and others, "Implement Port of Houston's Current Inventory and Harmonize the Remaining 8-county Shipping Inventory for TCEQ Modeling" Prepared for: Texas Commission on Environmental Quality, August 18, 2010.

Lindhjem, C.E. and Sturtz, T.M., "Development of Link-Level Spatial Allocation Methodology for Line-Haul Locomotive Emissions in Texas," Prepared for Texas Commission on Environmental Quality, August 15, 2009.

Lindhjem, C.E. 2008. "Evaluation Of Mobile Source Control Strategies For The Houston-Galveston-Brazoria State Implementation Plan," Draft Report, December 2008.

Lindhjem, C.E. 2008. "Emission Profiles for EPA SPECIATE Database, Part 2: EPAct FUELS." Prepared for EPA; Office of Transportation and Air Quality, Ann Arbor, MI. September.

Bar-Ilan, A., C. Lindhjem. 2008. "Port of Oakland Berthing Load Study." Prepared for the Port of Oakland, Oakland, CA. April.

Lindhjem, C.E. 2008. "Emission Profiles for EPA SPECIATE Database." Prepared for US Environmental Protection Agency. January.

Bar-Ilan, A., C. Chandler-Nogales, C. Lindhjem. 2007. "Massachusetts Diesel PM Emissions Reduction: Retrofit Technical Feasibility Analysis." Prepared for NESCAUM, Boston, MA. October.

Lindhjem, C.E. 2007. "Port of Oakland 2005 Seaport Air Emissions Inventory," Prepared for the Port of Oakland, Available online at http://www.portofoakland.com/environm/airEmissions.asp, August.

Lindhjem, C.E. 2007. "LADCO Nonroad Emission Inventory Project for Locomotive and Commercial Marine Emission Sources," Prepared for Lake Michigan Air Director Consortium by ENVIRON, February.

Lindhjem, C.E. 2006. "Los Angeles - Hobart Railyard Toxic Air Containment Emissions Inventory," "Commerce Eastern Railyard Toxic Air Containment Emissions," "Commerce – Mechanical Railyard Toxic Air Containment Emissions," "Wilmington Watson Railyard Toxic Air Containment Emissions," Stockton Railyard Toxic Air Containment Emissions," "Richmond Railyard Toxic Air Containment Emissions," "Barstow Railyard Toxic Air Containment Emissions," " San Diego Railyard Toxic Air Containment Emissions," Prepared for BNSF Railway, Available Online at http://www.arb.ca.gov/railyard/hra/hra.htm

Lindhjem, C.E. 2006. "Evaluation Of Mobile Source Control Strategies For The Houston-Galveston-Brazoria State Implementation Plan," Prepared for the Houston-Galveston Area Council, May.

Lindhjem, C.E. and Yarwood, G. 2004. "Humidity and Temperature Effects on On-road and Off-road Emissions and Ozone Formation (HARC Project H8B), Prepared for Houston Advanced Research Center (HARC), November.

Lindhjem, C.E., Shepard, S. 2004. "Development Work For Improved Heavy-Duty Vehicle Modeling Capability Data Mining – FHWA Datasets Phase II: Final Report," EPA Contract No. 68-C-02-022 Work Assignment No. 2-6 Prepared for Evelyn Sue Kimbrough, Atmospheric Protection Branch, Office of Research and Development, U.S. Environmental Protection Agency, September.

Lindhjem, C.E., Shepard, S., Pollack, A.K. 2004. "LADCO/MPCA Total Volume And Vehicle Classification Temporal Profiles," Prepared for Lake Michigan Air Directors Consortium, September.

Lindhjem, C.E. and Pollack, A.K. 2004. "Analysis of EPA's Draft Plan for Emissions Modeling in MOVES and MOVES GHG, Project E-68, Prepared for Coordinating Research Council, Inc., May.

ENVIRON. 2004. "Cold Ironing Cost Effectiveness," Prepared for the Port of Long Beach by ENVIRON International Corporation, March 30, 2004.

A. Pollack, R. Chi, C. Lindhjem, C. Tran, P. Chandraker. 2004. "Development of WRAP MOBILE Source Emission Inventories". Prepared for Western Governors' Association, Denver, Colorado. February.

A. Pollack, C. Lindhjem, T.E. Stoerckenius , C. Tran,G. Mansell, M. Jimenez, G. Wilson, and S. Coulter-Burke 2003. "CRC Project E-64 Evaluation of the U.S EPA MOBILE6 Highway Vehicle Emission Factor Model," Prepared for Coordinating Research Council, Inc., December.

Lindhjem, C.E., A. Pollack, R. Friesen, B. Sylte, D.Calkins, T. McGuire, D. Baldwin. 2003. "Potential Control Measures for Reducing Visibility-Related Pollutants in Maricopa County." Prepared for Arizona Department of Environmental Quality. December.

Lindhjem, C.E., A.K. Pollack, T.E. Stoerckenius, C. Tran, G. Mansell, M. Jimenez, G. Wilson, S.Coulter-Burke. 2003. "Evaluation of the US EPA MOBILE6 Highway Vehicle Emission Factory Model. December.

Lindhjem, C.E., A. Pollack, C. Tran, T. Stoeckenius, R. J. Downing, R. Schindler, E. Raisanen, D. Konopka, R. Sedlacek. 2003. "Maricopa County 2002 Comprehensive Emission Inventory; For the Cap and Trade Oversight Committee." Prepared for Arizona Department of Environmental Quality. October.

Lindhjem, C.E., Shepard, S. 2003. "Development Work For Improved Heavy-Duty Vehicle Modeling Capability, Phase I Final Report," EPA Contract No. 68-C-02-022 Work Assignment No. 1-7, Prepared for Evelyn Sue Kimbrough, Atmospheric Protection Branch, Office of Research and Development, U.S. Environmental Protection Agency, September.

Lindhjem, C.E., P. Chandraker. 2003. "Updates to Off-Road Mobile Emissions for East Texas." Prepared for the East Texas Council of Governments. September.

Lindhjem, C.E., M. Keinath, C. Tran. 2003. "Literature Survey of Mobile Source Air Toxic Emission Studies." Prepared for Office of Transportation and Air Quality; U.S. Environmental Protection Agency, September.

Lindhjem, C.E., S. Shepard. 2003. "Development Work for Improved Heavy-Duty Vehicle Modeling Capability." Prepared for Atmospheric Protection Branch; U.S. Environmental Protection Agency, September.

Lindhjem, C.E. 2002. "Improving Local Air Emissions Inventories From Nonroad Sources," Prepared for Lake Michigan Air Director Consortium, November.

Lindhjem, C.E., R.E. Morris, A.K. Pollack, G.E. Mansell, Ph.D., Y. Jia, G. Wilson. 2002. " Impact of Biodiesel Fuels on Air Quality and Human Health." Summary Report. Prepared for National Renewable Energy Laboratory. Golden, CO. November.

Lindhjem, C.E. 2002. "Review of EPA Draft Report: "Updating Fuel Economy Estimates in MOBILE6.3," Prepared for U.S. Environmental Protection Agency, Office of Transportation and Air Quality, June.

Lindhjem, C.E. 2002. "Commercial Marine Emission Inventory Development," Prepared for U.S. Environmental Protection Agency, Office of Transportation and Air Quality, April.

Lindhjem, C.E. and Stoeckenius, T., 2002. "On-Board Data Analysis and Collection for the New Generation Model, PR-CI-01-12239," Prepared for U.S. Environmental Protection Agency, Office of Transportation and Air Quality, Ann Arbor, MI. February.

Lindhjem, C.E., A.K. Pollack and R. Chi. 2001. "MOBILE6 Beta Testing." Prepared for U.S. Environmental Protection Agency, Office of Transportation and Air Quality, Ann Arbor, MI. January.

Lindhjem, C.E., C. Tran and A.K. Pollack. 2000. "Draft NONROAD Model Gasoline Consumption Estimates – Comparison with Other Data Sources." U.S. Environmental Protection Agency, October.

Lindhjem, C.E., 2000. "NONROAD and OFFROAD Models: Small (<25 hp) Spark-Ignition Engine Exhaust Emissions." U.S. Environmental Protection Agency, October.

Lindhjem, C.E., A.K. Pollack. 2000. "NONROAD and OFFROAD Models: Recreational Equipment Comparison." U.S. Environmental Protection Agency, October.

Lindhjem, C.E., (with others). 2000. "Study of Air Quality Programs in Clark County, Nevada," Pursuant to Senate Bill 432, Chapter 529, Statutes of Nevada 1999. Prepared for Legislative Counsel Bureau. September.

Lindhjem, C.E. 2000. "NONROAD and OFFROAD Models: Comparison of EPA and ARB Input Data for Emissions from Large Spark-Ignition Engines." U.S. Environmental Protection Agency, September.

Lindhjem, C.E. 2000. "Estimates of the Effect of the School Year on Vehicle Activity." Prepared for the Texas Natural Resource Conservation Commission. September.

Lindhjem, C.E., (with others). 2000. "On-Road Motor Vehicle NOx Emissions Reduction Measures for the Sacramento Metropolitan Area." Prepared for the Sacramento Council of Governments. August.

Lindhjem, C.E., (with others). 2000. “Houston-Galveston Area Marine Vessel Emission Inventory.” Prepared for the Port of Houston Authority and Texas Natural Resource Conservation Commission. June.

Lindhjem, C.E. 2000. “NONROAD and OFFROAD Models: Recreational Marine Comparison.” U.S. Environmental Protection Agency, February.

Lindhjem, C.E. 2000. “NONROAD and OFFROAD Models: Comparison of Modeling Input Data for Compression-Ignition Engines.” U.S. Environmental Protection Agency, January.

Lindhjem, C.E., (with others). 1999. “Control Measures for the Dallas-Ft. Worth.” Prepared for the North Central Texas Council of Governments. September.

Lindhjem, C.E., Coralee Cooper. 1999. “Heavy-Duty Nonroad Engine Activity and Emissions in the Northeast.” Northeast States for Coordinated Air Use Management, August.

Lindhjem, C.E., A.K. Pollack. 1999. “TNRCC Construction Equipment Emissions Project.” Prepared for the Texas Natural Resources Conservation Commission. February.

Lindhjem, C.E. 1998. “ARB OFFROAD and EPA NONROAD Models: Comparison of Model Structures.” U.S. Environmental Protection Agency, December.

Lindhjem, C.E., A.K. Pollack. 1998. “Comparison of EPA NONROAD and ARB OFFROAD models: Population Estimates and Allocation to Counties for Small (<25 hp) Spark Ignition (SI) Engines.” U.S. Environmental Protection Agency, December.

Lindhjem, C.E., M. Beardsley, C. Harris. 1998. “Exhaust Emission Factors for Nonroad Engine Modeling--Spark Ignition.” Report No. NR-010, Documentation Supporting EPA’s Draft Release of the NONROAD Model. February.

Lindhjem, C.E., M. Beardsley. 1998. “Exhaust Emission Factors for Nonroad Engine Modeling--Compression-Ignition.” Report No. NR-009, Documentation Supporting EPA’s Draft Release of the NONROAD Model. February.

Lindhjem, C.E., 1997. “Nonroad Engine Population Estimates.” Report No. NR-006, Documentation Supporting EPA’s Draft Release of the NONROAD Model. December.

Lindhjem, C.E., M. Beardsley. 1997. “Average Life, Annual Activity, and Load Factor Values for Nonroad Engine Emissions Modeling.” EPA/OMS Report NR-005, Documentation Supporting EPA’s Draft Release of the NONROAD Model. December.

Lindhjem, C.E., 1997. “Exhaust Emission Effects of Fuel Sulfur and Oxygen on Gasoline Nonroad Engines.” Report No. NR-003, Documentation Supporting EPA’s Draft Release of the NONROAD Model. November.

Lindhjem, C.E., 1997. “Conversion Factors for Hydrocarbon Emission Components.” Report No. NR-002, Documentation Supporting EPA’s Draft Release of the NONROAD Model. November.

Lindhjem, C.E., T. Jackson. 1997. “Update on Heavy-Duty Emission Levels (Model Years 1988-2004+) for Use in MOBILE6.” Draft Report M6HDE.001. October.

Lindhjem, C.E., 1997. “Nonroad Mobile Emissions Modeling.” Air & Waste Management Conference on Emission Inventory: Planning for the Future, RTP, NC. October.

Lindhjem, C.E., 1997. “Heavy-Duty Diesel Emissions Modeling.” Air & Waste Management Conference on Emission Inventory: Planning for the Future, RTP, NC. October.

Lindhjem, C.E. 1997. “Emission Rates from Highway and Nonroad Diesel Engines.” World Car Conference, Riverside, California. January.

Lindhjem, C.E., 1995. “Heavy-duty Engine Testing Report; Emissions from Idling Truck Engines.” EPA Technical Report. May.

Lindhjem, C.E., 1994. “The Effect of Fuel Reformulation on 4-stroke Lawn and Garden Engines.” EPA Technical Report. July.

Lindhjem, C.E., 1993. “The Exhaust VOC Emission Inventory by Vehicle Emitter Class Following Implementation of an Enhanced Inspection and Maintenance (I/M) Program.” EPA Technical Report (for use in EPA's rulemaking on Reformulated Gasoline using the Complex Model). June.

Lindhjem, C.E., 1993. “Speciation of Evaporative Emissions.” EPA Technical Report. April.
Lindhjem, C.E., 1992. “Exhaust, Evaporative, and Running Loss Speciation.” EPA Technical Report. July.
Lindhjem, C.E., 1992. “Vapor Pressure at 130^{o}F as the Parameter of choice for Running Loss and Hot Soak Emissions Correlation.” EPA Technical Report. April.
Lindhjem, C.E., 1992. “Effect of Oxygenates on Emissions.” EPA Technical Report (for use in EPA's Rulemaking on Reformulated Gasoline using the Simple Model). January.
Lindhjem, C.E., 1991. “Multiple Day Diurnal, Purge Failure, and Pressure Failure Evaporative Emissions.” EPA Technical Report (for use in EPA's MOBILE emission factors model for mobile source emission estimations). November.

Brian H. West, P.E.
Deputy Director of Fuels, Engines, and Emissions Research Center
Oak Ridge National Laboratory
NTRC Building, 2360 Cherahala Blvd
Knoxville, TN 37932
865-946-1231

EDUCATION

M. S. Mechanical Engineering, Virginia Tech, 1988
B. S. Mechanical Engineering, Clemson University, 1986

PROFESSIONAL HISTORY

Oak Ridge National Laboratory, Oak Ridge, TN, 5/88 – present

- Development Engineer and Deputy Director of Fuels, Engines, and Emissions Research Center
- Design, planning, management, and execution of projects to evaluate advanced engines, catalysts and emission control systems, materials, and alternative fuels.
- Leader for ORNL vehicle research activities, including procedure/protocol development and assessment and characterization of alternative fuels, advanced diesel emissions control, full electric prototype USPS vehicles, unique European vehicles, etc.
- Leader for Department of Energy (DOE) / Coordinating Research Council (CRC) test program to investigate the effects of mid-level ethanol blends on vehicles.
- Co-chair of NOx adsorber technical workgroup for Advanced Petroleum-Based Fuels - Diesel Emission Control Program, an industry-government cooperative program.
- Leader for Department of Transportation program to develop modal emissions and fuel consumption models for traffic simulations

MPR Associates, Incorporated, Alexandria, VA, 2/98 - 8/98

Senior Engineer at this Engineering Consulting Firm. Developed engine governor troubleshooting manual for emergency diesel generators operated by nuclear power plants. Conducted electrical power plant assessments in Brazil, including coal-fired, oil-fired, and hydroelectric plants.

Hoechst Fibers Industries, Spartanburg, SC, 1/84 - 8/85

Co-op Engineer at this synthetic fiber producer for 3 semesters during undergraduate study at Clemson University. Assisted senior engineering staff in maintenance and development engineering in the plant.

Motorola Portable Products, Fort Lauderdale, FL, 1/83 - 5/83

Co-op Engineer at this pager development and manufacturing facility. Assisted Mechanical Laboratory manager and other engineers in testing components and maintaining equipment for shock, vibration, and strength testing.

Professional Activities, Honors, Awards

- DOE R&D Award for Managing Mid-Level Ethanol Blends Vehicle aging program (2010 DOE Merit Review)
- Member of Society of Automotive Engineers (23 years).
- Recipient of SAE Lloyd L. Withrow Distinguished Speaker Award (April 2009).
- Two patents
- Invited Technical Panelist for DOE Merit Review
- Licensed Professional Engineer in Tennessee

Selected Publications

West, Brian H., C. Scott Sluder, Keith E. Knoll, John E. Orban, and Jingyu Feng, "Intermediate Ethanol Blends Catalyst Durability Program," ORNL/TM-2011/234.

Sluder, C. Scott, and Brian West, "Limitations and Recommended Practice In the Use of Compression and Leak-Down Tests to Monitor Gradual Engine Degradation," SAE Int. J. Engines, Vol. 4, Issue 3, December 2011.

Sluder, C. Scott, and Brian H. West, "NMOG Emissions Characterizations and Estimation for Vehicles Using Ethanol-Blended Fuels," ORNL/TM-2011/461, October 2011.

Knoll, Keith, Brian West, Shean Huff, John Thomas, John Orban, Cynthia Cooper, "Effects of Mid-Level Ethanol Blends on Conventional Vehicle Emissions," Society of Automotive Engineers Paper 2009-01-2723, November 2009.

Knoll, Keith, Brian West, Wendy Clark, Ronald Graves, John Orban, Steve Przesmitzki, Timothy Theiss, "Effects of Intermediate Ethanol Blends on Legacy Vehicles and Small Non-Road Engines, Report 1 – Updated," ORNL/TM-2008/117, February 2009.

Norman, Kevin, Shean Huff, and Brian West, "Effect of Intake Air Filter Condition on Vehicle Fuel Economy," ORNL/TM-2009/021, February 2009.

Parks, James E., Brian West, and Shean Huff, "Characterization of Lean NOx Trap Catalysts with In-Cylinder Regeneration Strategies," SAE Paper Number 2008-01-0448, April 2008

West, Brian H., Alberto J. Lopez, Timothy J. Theiss, Ronald L. Graves, John Storey, and Sam Lewis "Fuel Economy and Emissions of the Ethanol-Optimized Saab 9-5 BioPower," Society of Automotive Engineers Paper 2007-01-3994, October 2007.

Bechtold, R., J. F. Thomas, S. P. Huff, J. P. Szybist, T. J. Theiss, B. H. West, M. Goodman, T. A. Timbario, "Technical Issues Associated with the Use of Intermediate Ethanol Blends (>E10) in the U.S. Legacy Fleet: Assessment of Prior Studies," ORNL/TM-2007/37, August 2007.

Swartz, Matthew M., Shean Huff, James Parks, and Brian West, "Intra-Catalyst Reductant Chemistry and NOx Conversion of Diesel Lean NOx Traps at Various Stages of Sulfur Loading," Society of Automotive Engineers Paper 2006-01-3423, October 2006.

West, Brian H., Shean Huff, James E. Parks, Matt Swartz, and Ronald L. Graves, "In-Cylinder Production of Hydrogen During Net-Lean Diesel Operation," Society of Automotive Engineers Paper 2006-01-0212, April 2006

Huff, Shean, Brian West, James E. Parks, Matt Swartz, Johney Green, and Ronald Graves, "In-Cylinder Regeneration of Lean NOx Trap Catalysts Using Low Temperature Combustion," Society of Automotive Engineers Paper 2006-01-1416, April 2006

Parks, James E., Shean Huff, Josh Pihl, Jae-Soon Choi, and Brian H. West, "Nitrogen Selectivity in Lean NOx Trap Catalysis With Diesel Engine In-Cylinder Regeneration," Society of Automotive Engineers Paper 2005-01-3876, October 2005.

Witze, Peter O., Shean P. Huff, John M. Storey, Brian H. West, "Time-Resolved, Laser Induced Incandescence Measurements of Particulate Emissions During Enrichment for Diesel Lean NOx Trap Regeneration," Society of Automotive Engineers Paper 2005-01-0186, April 2005.

West, Brian H., Shean P. Huff, James E. Parks, Jae-Soon Choi, William P. Partridge, Sam A. Lewis, John M. Storey, "Assessing Reductant Chemistry During In-Cylinder Regeneration of Diesel Lean NOx Traps," Society of Automotive Engineers Paper 2004-01-3023, October 2004.

Sluder, C. Scott, and Brian H. West, "Effects of Regeneration Conditions on NOx Adsorber Performance," Society of Automotive Engineers Paper 2002-01-2876, October 2002.

Sluder, C. Scott, and Brian H. West, "Performance of a NOx Adsorber and Catalyzed Particle Filter System on a Light-Duty Diesel Vehicle," Society of Automotive Engineers Paper 2001-01-1933, May 2001.

West, Brian H., and C. Scott Sluder, "NOx adsorber performance in a Light-Duty Diesel Vehicle," Society of Automotive Engineers Paper 2000-01-2912, October 2000.

Sluder, C. Scott, and Brian H. West, "Catalyzed Diesel Particulate Filter Performance in a Light-Duty Vehicle," Society of Automotive Engineers Paper 2000-01-2848, October 2000.

West, B. H., R. N. McGill, J. W. Hodgson, C. S. Sluder, D. E. Smith "Development of Data-Based, Light-Duty Modal Emissions and Fuel Consumption Models," Society of Automotive Engineers Paper 972910, October 1997.

West, B. H., R. N. McGill, C. S. Sluder, and J. W. Hodgson, “Data-Based Modal Emissions and Fuel Consumption Models for Transportation Assessments,” Air and Waste Management Association, *Proceedings of Emission Inventory: Planning for the Future*, October 1997.

West, B. H., J. M. Storey, S. A. Lewis, G. L. Devault, J. B. Green, C. S. Sluder, J. W. Hodgson, and B. L. Moore, “Experimental and Numerical Assessment of On-Road Diesel and Biodiesel Emissions,” *Proceedings of 1997 Diesel Engine Emissions Reduction Workshop*, July 1997.

West, B., J. B. Green, *Effects of Piston Surface Treatments on Performance and Emissions of a Methanol-Fueled, Direct-Injection, Stratified-Charge Engine*, National Renewable Energy Laboratory Report NREL/TP-425-6161, July 1994.

West, B. H., R. N. McGill, S. L. Hillis, and J. W. Hodgson, *Federal Methanol Fleet Project Final Report*, Oak Ridge National Laboratory Report ORNL/TM-12278, March 1993.

West, B. H., and R. N. McGill, "Oil Performance in a Methanol-Fueled Vehicle Used in Severe Short-Trip Service," Society of Automotive Engineers Paper 922298, October 1992.

McGill, R. N., R. L. Graves, B. H. West, and J. W. Hodgson, "The Federal Methanol Fleet - Summary of Technical Data," Published in proceedings of *IX International Symposium on Alcohol Fuels*, Firenze, Italy, November 12-15, 1991.

West, B. H., R. R. Judkins, R. L. Graves, C. S. Daw, R. P. Krishnan, A. L. Compere, and W. L. Griffith, "Mild-Gasification Product Characterization--1991 Results," in *Proceedings of the Eleventh Annual Gasification and Gas Stream Cleanup Systems Contractors Review Meeting*, Vol. 1, DOE/METC-91/6123, Morgantown, WV, August 1991.

McGill, R. N., and B. H. West, "The Federal Methanol Fleet -A Demonstration of Methanol Vehicle Technology," Published in *Proceedings of Technical and Business Exposition and Symposium*, Huntsville, AL, May 14-15, 1991.

Brinkman, C. R., G. M. Begun, R. L. Graves, W. K. Kahl, K. C. Liu, and B. H. West, *Influence of Diesel Engine Exposure on the Rupture Strength of Silicon Nitride and Partially Stabilized Zirconia*, Oak Ridge National Laboratory Report ORNL/TM-6612, June 1990.

Daw, C. S., R. L. Graves, R. P. Krishnan, B. H. West, A. L. Compere, and W. L Griffith, "Mild-Gasification Product Characterization -- 1990 Results," in *Proceedings of the Tenth Annual Gasification and Gas Stream Cleanup Systems Contractors Review Meeting*, Vol. II, DOE/METC-90/6115, Morgantown, WV, August 1990.

West, B. H., A. L. Compere, W. L. Griffith, "High-Alcohol Microemulsion Fuel Performance in a Diesel Engine," Society of Automotive Engineers Paper 902101, October 1990.

Graves, R. L., C. S. Daw, R. P. Krishnan, and B. H. West, "Mild-Gasification Product Characterization," in *Proceedings of the Ninth Annual Gasification and Gas Stream Cleanup Systems Contractors Review Meeting*, Vol. II, DOE/METC-89/6107, Morgantown, WV, June 1989.

Nelson, D. J., and B. H. West, "Fully Developed Natural and Mixed Convection in Inclined Rectangular Ducts," American Society of Mechanical Engineers paper published in *Proceedings of 1988 Winter Annual Meeting*, December 1988.

Appendix B: Conflict of Interest Statements

Conflict of Interest and Bias for Peer Review

Background

Identification and management of potential conflict of interest (COI) and bias issues are vital to the successes and credibility of any peer review consisting of external experts. The questionnaire that follows is consistent with EPA guidance concerning peer reviews.[1]

Definitions

Experts in a particular field will, in many cases, have existing opinions concerning the subject of the peer review. These opinions may be considered bias, but are not necessarily conflicts of interest.

Bias: For a peer review, means a predisposition towards the subject matter to be discussed that could influence the candidate's viewpoint.

Examples of bias would be situations in which a candidate:

1. Has previously expressed a position on the subject(s) under consideration by the panel; or

2. Is affiliated with an industry, governmental, public interest, or other group which has expressed a position concerning the subject(s) under consideration by the panel.

Conflict of Interest: For a peer review, as defined by the National Academy of Sciences,[2] includes any of the following:

1. Affiliation with an organization with financial ties directly related to the outcome;

2. Direct personal/financial investments in the sponsoring organization or related to the subject; or

3. Direct involvement in the documents submitted to the peer review panel... that could impair the individual's objectivity or create an unfair competitive advantage for the individual or organization.

Policy and Process

- Candidates with COI, as defined above, will not be eligible for membership on those panels where their conflicts apply.

---

[1] U.S. EPA (2009). Science Policy Council Peer Review Handbook. OMB (2004). Final Information Quality Bulletin for Peer Review.

[2] NAS (2003). "Policy and Procedures on Committee Composition and Balance and Conflict or Interest for Committees Used in the Development of Reports" (www.nationalacademies.org/coi).

- In general, candidates with bias, as defined above, on a particular issue will be eligible for all panel memberships; however, extreme biases, such as those likely to impair a candidate's ability to contribute to meaningful scientific discourse, will disqualify a candidate.

- Ideally, the composition of each panel will reflect a range of bias for a particular subject, striving for balance.

- Candidates who meet scientific qualifications and other eligibility criteria will be asked to provide written disclosure through a confidential questionnaire of all potential COI and bias issues during the candidate identification and selection process.

- Candidates should be prepared, as necessary, to discuss potential COI and bias issues.

- All bias issues related to selected panelists will be disclosed in writing in the final peer review record.

Conflict of Interest and Bias Questionnaire

EPAct Study Analysis Peer Review

Instructions to Candidate Reviewers

1. Please check YES/NO/DON'T KNOW in response to each question.

2. If your answer is YES or DON'T KNOW, please provide a brief explanation of the circumstances.

3. Please make a reasonable effort to answer accurately each question. For example, to the extent a question applies to individuals (or entities) other than you (e.g., spouse, dependents, or their employers), you should make a reasonable inquiry, such as emailing the questions to such individuals/entities in an effort to obtain information necessary to accurately answer the questions.

Questions

1. Are you (or your spouse/partner or dependents) or your current employer, an author, contributor, or an earlier reviewer of the document(s) being reviewed by this panel?

   YES___ NO_X_ DON'T KNOW___

2. Do you (or you spouse/partner or dependents) or your current employer have current plans to conduct or seek work related to the subject of this peer review following the completion of this peer review panel?

   YES___ NO_X_ DON'T KNOW___

3. Do you (or your spouse/partner or dependents) or your current employer have any known financial stake in the outcome of the review *(e.g.,* investment interest in a business related to the subject of peer review)?

   YES___ NO_X_ DON'T KNOW___

4. Have you (or your spouse/partner or dependents) or your current employer commented, reviewed, testified, published, made public statements, or taken positions regarding the subject of this peer review?

   YES___ NO_X_ DON'T KNOW___

5. Do you hold personal values or beliefs that would preclude you from conducting an objective, scientific evaluation of the subject of the review?

   YES___ NO_X_ DON'T KNOW___

6. Do you know of any reason that you might be unable to provide impartial advice or comments on the subject review of the panel?

   YES___ NO_X_ DON'T KNOW___

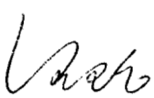
7. Are you aware of any other factors that may create potential conflict of interest or bias issues for you as a member of the panel?

   YES___ NO_X_ DON'T KNOW___

Acknowledgment

I declare that the disclosed information is true and accurate to the best of my knowledge, and that no real, potential, or apparent conflict of interest or bias is known to me except as disclosed. I further declare that I have made reasonable effort and inquiry to obtain the information needed to answer the questions truthfully, and accurately. I agree to inform SRA promptly of any change in circumstances that would require me to revise the answers that I have provided.

Xuming He________________
Name

 ____________________ ______11/15/11________
Signature Date

Conflict of Interest and Bias Questionnaire

EPAct Study Analysis Peer Review

Instructions to Candidate Reviewers

1. Please check YES/NO/DON'T KNOW in response to each question.

2. If your answer is YES or DON'T KNOW, please provide a brief explanation of the circumstances.

3. Please make a reasonable effort to answer accurately each question. For example, to the extent a question applies to individuals (or entities) other than you (e.g., spouse, dependents, or their employers), you should make a reasonable inquiry, such as emailing the questions to such individuals/entities in an effort to obtain information necessary to accurately answer the questions.

Questions

1. Are you (or your spouse/partner or dependents) or your current employer, an author, contributor, or an earlier reviewer of the document(s) being reviewed by this panel?

   YES___ NO___ DON'T KNOW_X_

   [Chris Lindhjem may have investigated this data on behalf of EPA to develop speciation (chemical composition) estimates for emissions from these fuels. The description of this work primarily focuses on regulated emissions, so speciation would not be an element of this review.]
2. Do you (or you spouse/partner or dependents) or your current employer have current plans to conduct or seek work related to the subject of this peer review following the completion of this peer review panel?

   YES___ NO___ DON'T KNOW_X_

   [ENVIRON has had a business relationship with an ethanol producer reviewing emissions data related to fuel certification.]
3. Do you (or your spouse/partner or dependents) or your current employer have any known financial stake in the outcome of the review *(e.g.,* investment interest in a business related to the subject of peer review)?

   YES___ NO_X_ DON'T KNOW___

4. Have you (or your spouse/partner or dependents) or your current employer commented, reviewed, testified, published, made public statements, or taken positions regarding the subject of this peer review?

   YES___ NO_X_ DON'T KNOW___

5. Do you hold personal values or beliefs that would preclude you from conducting an objective, scientific evaluation of the subject of the review?

   YES___ NO_X_ DON'T KNOW___

6. Do you know of any reason that you might be unable to provide impartial advice or comments on the subject review of the panel?

   YES___ NO_X_ DON'T KNOW___

7. Are you aware of any other factors that may create potential conflict of interest or bias issues for you as a member of the panel?

   YES___ NO_X_ DON'T KNOW___

Acknowledgment

I declare that the disclosed information is true and accurate to the best of my knowledge, and that no real, potential, or apparent conflict of interest or bias is known to me except as disclosed. I further declare that I have made reasonable effort and inquiry to obtain the information needed to answer the questions truthfully, and accurately. I agree to inform SRA promptly of any change in circumstances that would require me to revise the answers that I have provided.

Chris Lindhjem
Name

Signature

January 9, 2012
Date

Conflict of Interest and Bias Questionnaire

EPAct Study Analysis Peer Review

Instructions to Candidate Reviewers

1. Please check YES/NO/DON'T KNOW in response to each question.

2. If your answer is YES or DON'T KNOW, please provide a brief explanation of the circumstances.

3. Please make a reasonable effort to answer accurately each question. For example, to the extent a question applies to individuals (or entities) other than you (e.g., spouse, dependents, or their employers), you should make a reasonable inquiry, such as emailing the questions to such individuals/entities in an effort to obtain information necessary to accurately answer the questions.

Questions

1. Are you (or your spouse/partner or dependents) or your current employer, an author, contributor, or an earlier reviewer of the document(s) being reviewed by this panel?

   YES___ NO_X_ DON'T KNOW___

2. Do you (or you spouse/partner or dependents) or your current employer have current plans to conduct or seek work related to the subject of this peer review following the completion of this peer review panel?

   YES___ NO___ DON'T KNOW_X_

   In 23 years at QRNL I have worked on a wide range of fuels, engine, and emissions control technologies, conducting experiments in engine labs, vehicle dynamometer labs, or managing subcontracts at contractor facilities such as SwRI or TRC. I presume that I am a candidate reviewer because of this experience. Future work at ORNL could very well be related to the subject of the EPAct study, just as much of my past work experience has been; however, I do not currently have plans that are related to the EPAct study.

3. Do you (or your spouse/partner or dependents) or your current employer have any known financial stake in the outcome of the review *(e.g.*, investment interest in a business related to the subject of peer review)?

   YES___ NO_X_ DON'T KNOW___

4. Have you (or your spouse/partner or dependents) or your current employer commented, reviewed, testified, published, made public statements, or taken positions regarding the subject of this peer review?

   YES___ NO___ DON'T KNOW_X_

   I endorsed using DOE funds to augment the EPA funding for the subject test program. I managed a large DOE program (DOE V4 program) from 2008-2011 in which over 80 vehicles were aged and emissions tested with ethanol blends (at SwRI and TRC). Emissions measurements in the V4 program were largely limited to criteria pollutants. Several national lab colleagues and I saw value in the detailed exhaust speciation work to be done at SwRI under the EPAct program, and endorsed having DOE leverage that effort. DOE funds (through NREL contract with SwRI) supported part of the EPAct (DOE V2) project to allow a wider range of fuels, including ethanol blends (of particular interest to DOE). I shared my position on the matter with DOE sponsors and NREL, ORNL, and EPA colleagues in meetings in 2008. I have not testified or published any statements to this effect

5. Do you hold personal values or beliefs that would preclude you from conducting an objective, scientific evaluation of the subject of the review?

   YES___ NO_X_ DON'T KNOW___

6. Do you know of any reason that you might be unable to provide impartial advice or comments on the subject review of the panel?

   YES___ NO_X_ DON'T KNOW___

7. Are you aware of any other factors that may create potential conflict of interest or bias issues for you as a member of the panel?

   YES___ NO_X_ DON'T KNOW___

Acknowledgment

I declare that the disclosed information is true and accurate to the best of my knowledge, and that no real, potential, or apparent conflict of interest or bias is known to me except as disclosed. I further declare that I have made reasonable effort and inquiry to obtain the information needed to answer the questions truthfully, and accurately. I agree to inform SRA promptly of any change in circumstances that would require me to revise the answers that I have provided.

Brian West________________
Name

[signature]________________ ___________
Signature Date

Appendix C: Peer Review Charge

Charge to Peer Reviewers of *Assessing the Effect of Five Gasoline Properties on Exhaust Emissions from Light-Duty Vehicles certified to Tier-2 Standards (EPAct/V2/E-89: Phase 3)*

Past fuel effects models, such as the EPA Predictive Model and EPA's Complex Model, were based on data collected using 1990s-technology cars and trucks meeting Tier 0 and Tier 1 vehicle emission standards at pollutant levels an order of magnitude higher than compliance levels for current Tier2 vehicles. With the current on-highway fleet turning out much lower emitting vehicles than in past years, the U.S. Congress, the EPA and a variety of stakeholders are interested in generating an updated data set for fuel effects models to guide vehicle emissions policy going forward. Further, Section 1506 of the Energy Policy Act (EPAct) of 2005 directed EPA to generate a more current fuel effects model representing future gasoline vehicle fleet emissions. Statutory requirements such as the one outlined in Section 209 of the Energy Independence and Security Act (EISA) of 2007, an anti-backsliding assessment of the effects of increased renewable fuels use on air quality, are likewise dependent upon the output of updated fuel effects models.

This report describes the analysis and modeling of data collected in Phase 3 of the EPAct/V2/E-89 light duty gasoline vehicle fuel effects study (referred to here as "the EPAct program", or "Phase 3 of the EPAct program"). This study examined the exhaust emission impacts on vehicles of changing levels of five fuel properties (ethanol, T50, T90, aromatics, and RVP (specified as DVPE)) in a matrix of 27 gasoline fuels, arranged according to a partial factorial design optimized for selected interactions of interest. The range in the levels of fuel properties tested is comparable to those seen in current market fuels and in potential mid-level ethanol blended fuels. The test vehicle fleet consisted of 15 new 2008 light-duty cars and trucks, selected from among high sales makes and models, to provide a representative sample of a fleet of properly-operating vehicles meeting the U.S. Federal Tier 2 vehicle emission standards. Given the relatively low level of emissions from these vehicles, a number of design and procedural steps were undertaken to minimize the impacts of measurement variability and other artifacts on data quality.

Vehicle emission test data collected in the program included concentrations and rates for typical regulated pollutants by bag and by second-by-second collection, plus data for speciated emissions for a subset of vehicle emission tests and bags. Complete data was generated for 926 tests, with 30 additional tests containing valid measurements for regulated emissions. The analysis of this data focuses on producing reduced mixed-effect models by emission bag for the pollutants of interest (nitrogen oxides, total hydrocarbon, non-methane organic gases and carbon monoxide) after screening for outliers and other data quality issues. Fuel effects models were produced using the target set of six interactive terms, for which the fuel matrix was optimized, plus a set of terms generated by using all possible interactive terms (eleven) and evaluating fit parameters for both. A conservative approach was taken on outliers, such that abnormal statistical parameters were generally not sufficient to remove data without an underlying measurement or procedural problem. Relatively few outliers were removed based on having an unrealistically high value (only four measurements, all from the particulate mass emission dataset), but portions of the nitrogen oxides data were removed after data quality review suggested some very low-level measurements from certain vehicles were suspected to contain a large amount of measurement error. Another issue encountered during analysis was treatment of non-detects (zero values), which were addressed via use of Tobit regressions (i.e., the PROC LIFEREG procedure in SAS).

The primary objective of this program is to generate data to produce models with good predictive abilities for vehicles and fuels beyond those specifically tested in this program. Thus, minimizing retention of terms that are based largely on measurement or design artifact or are otherwise overfit to these vehicles and fuels is a concern.

EPA is seeking your expert opinion on the appropriateness of the statistical techniques described in the *EPAct Study Analysis* and their appropriateness in the context of any data accuracy/quality issues. Given your knowledge and understanding of the accuracy of the test measurements, has EPA chosen the most useful means of statistical analysis? Model validation using external datasets and crossvalidation (testing predictive ability using portions of data omitted during model fitting) are only addressed briefly here and remain as possible subjects of ongoing report and comment and future study.

Some specific areas of focus include the following:

1. Was the process of imputation of NMOG/NMHC results for tests/bags with missing speciation data reasonable and statistically sound? (Section X.X)
2. Was the decision to remove very low emitting and influential vehicles from the NOx and NMOG analyses reasonable? (Sections 5.5 and 6.1)
3. Please comment on the use of the “design set” of 11 terms as the basis of the final models, versus allowing model to fit all 17 terms, including the adequacy of the justification of this decision. (Section 7)
4. Comment on the use of the Tobit regression (SAS PROC LIFEREG) for modeling datasets with large numbers of censored values. (Section 5.3)
5. Comment on the decision to independently model the linear terms and interactions between fuel blends as presented in the final results?
6. Please comment on the methods used to select reduced models. (Section 5)

In addition to addressing these issues, EPA encourages you to best apply your particular area(s) of expertise to review the overall study.

In your comments you should distinguish between recommendations for clearly defined improvements that can be readily made based on data or literature reasonably available to EPA and improvements that are more exploratory or dependent on information not readily available to EPA. Your comments should be sufficiently clear and detailed to allow readers to thoroughly understand their relevance to the *EPAct Study Analysis*. ***<u>Please deliver your final written comments to SRA International by Friday, January 20, 2012.</u>***

All materials provided to you as well as your review comments should be treated as confidential, and should neither be released nor discussed with others outside of the review panel. Once EPA has made its report and supporting documentation public, the Agency will notify you that you may release or discuss the peer review materials and your review comments with others.

Should you have questions about what is required in order to complete this review or need additional background material, please contact Brian Menard at SRA (Brian_Menard@sra.com) or (434-817-4133).

Appendix D: Reviews

Comments on the EPAct Study Analysis

Xuming He

The report under review describes the analysis and modeling of data collected in Phase 3 of the EPAct/V2/E-89 light duty gasoline vehicle fuel effects study. The study examined the exhaust emission impacts on vehicles of changing levels of five fuel properties (ethanol, T50, T90, aromatics, and RVP (specified as DVPE)) in a matrix of 27 gasoline fuels, arranged according to a partial factorial design optimized for selected interactions of interest. A number of design and procedural steps were undertaken in this study to reduce the impacts of measurement bias and variability on data quality and statistical analysis. Overall, I found the study well designed and carefully analyzed with generally accepted modern statistical tools. My comments will focus on the appropriateness of the data processing and statistical techniques described in the report, with the purpose of improving the EPAct Study Analysis. The item numbers used below (1 – 6) correspond to the prompts given in the email of December 22 from Brian Menard. Some additional comments and suggestions are given at the end of this review.

1. Imputation of NMOG/NMHC results for tests/bags with missing speciation data

Due to the speciation schedule described in Section 2.2 of the report, most tests in the dataset do not have alcohol and carbonyl measurements for bags 2 and 3. As NMOG and NMHC are calculated emission results that use speciation data, they could not be computed for the portions of the dataset without speciation. The study used imputation based on an alternate measure of hydrocarbon emissions to fill in the missing values. Linear location-scale-type models were used to imputation with special substitution of value zero for small NMOG. The report made a convincing case that the models used for imputation fit the data well, and should result in small errors and variability due to imputation.

It seems that the imputed values were deterministic in the study given the variables x_{NMHC} in Equations 8-11. If so, statistical variability could be under-reported in the subsequent studies. One approach to recommend here is to use multiple imputations to account for the variability. Based on the descriptions in the report, I do not think that the additional variability due to imputation would be a significant factor, but I would prefer to see a more explicit discussion and examination of this issue in the study.

2. Decision to remove very low emitting and influential vehicles from the NOx and NMOG analyses;

Modifying or removing outliers and influential observations could raise questions about the validity of a study, especially when ad hoc decisions are made after the data are collected and examined. The report paid serious attention to data quality, and described how outliers and influential observations were identified.

I believe that the decision to use left-censored models in the study is appropriate. This allows linear models to remain valid for the data with left end points. Such practices have been used in the statistics and econometrics literature. There are, however, several minor issues to deal with.

(a) In Section 5.2, censored measurements were replaced by the minimum positive value measured for the emission and bag. This substitution has the potential to lower the variance estimate of the statistical models, unless censoring is taken into account in the variance estimates. If the error variance estimates are deflated, we would see higher “studentized residuals”, resulting in false positives in outlier detection. The issue needs to be carefully examined.

(b) At the end of Section 5.2, it was mentioned that Run 6281 in Bag 3 was removed even though it was not flagged as influential. The specific reason for this decision was lacking in the report.

(c) In the analysis of Section 5.3.1, a distinction was made between light censoring and severe censoring. In the case of severe censoring, the Tobit regression was used in the data analysis. Otherwise, the censored values were substituted. I do not see good reasons for handling the two scenarios differently. Why not use the Tobit regression in all cases? The current practice would raise a question about the stability and sensitivity of the results if one more or one fewer data point is censored.

(d) In Section 5.3.2, both BIC (model selection criterion) and likelihood ratio tests were described and used. Although both approaches are valid and useful, they have different goals in mind. Model selection based on BIC is to choose models, treating all competing models equally. The chi-square tests have a null hypothesis in mind, where the null hypothesis refers to the smaller models in the present analysis. Test decisions are designed to protect the null hypothesis, so the competing models are not treated equally. When both approaches are used, one needs to be clear how they work together, and what are to be achieved. I do not imply that anything has gone wrong here, but this part of the analysis needs to be made clearer as to why one cannot simply use BIC.

3, 5 and 6. Use of the “design set” of 11 terms as the basis of the final models, versus allowing model to fit all 17 terms, including the adequacy of the justification of this decision; the decision to independently model the linear terms and interactions between fuel blends as presented in the final results; the methods used to select reduced models.

Table 36 report correlation coefficients between the linear-effects and the additional terms (interactions). I would suggest including the canonical correlation between the set of linear-effects and the set of interactions. This would assess linearity beyond pairwise correlations.
The considerations and justifications on finding final models (Section 7.2) were reasonable and well thought-out. Because some subjectivity was involved in the final model selections, it is hard for me to tell whether each detailed review and scrutiny reported in Section 7.3 is “optimal”. On the other hand, there is no single model that is likely to be *the best*. In reality, it is generally the case that several models are (almost) equally good given the limited amount of data, and the final selection can be made with some subjectivity. The analyses given in Section 7.3 appear reasonable.

4. Use of the Tobit regression (SAS PROC LIFEREG) for modeling datasets with large numbers of censored values

I think that the use of left-censored models and the Tobit regression is appropriate. My only question here is why the same approach is not taken for cases with light censoring. See my comments under Item #2.

Additional suggestions

I have some additional suggestions, some of which might be useful for future studies. First, if a similar study is planned in the future, it would be better to construct specific criteria for removing outliers prior to data collection. This would eliminate questions about biased interference in the data processing stage. Second, when imputation is used, more careful procedures should be in place to account for variability due to imputation. Multiple imputation is a common approach to take. Third, some sensitivity analysis using robust statistical methods can be performed to understand the effects of outlying/influential points and their handling on the final analysis. Most statistical techniques, including linear mixed models and the Tobit regression used in this study, are based on the assumption of Gaussian errors. Robust statistical methods can help us understand the impact of non-Gaussian errors.

January 20, 2012

MEMORANDUM

To: Brian Menard
From: Chris Lindhjem, ENVIRON International
Subject: Review of "Assessing the Effect of Five Gasoline Properties on Exhaust Emissions from Light-Duty Vehicles certified to Tier-2 Standards (EPAct/V2/E-89: Phase 3) Part II: Data Analysis and Model Development" DRAFT REPORT December, 2011

INTRODUCTION

The report, "Assessing the Effect of Five Gasoline Properties on Exhaust Emissions from Light-Duty Vehicles certified to Tier-2 Standards (EPAct/V2/E-89: Phase 3) Part II: Data Analysis and Model Development DRAFT REPORT" presents a well-documented approach to estimating the effect that gasoline fuel properties have on late model vehicle emissions. In general, the approach to evaluating the data is equal to or more robust than previous efforts, such as the complex and predictive models or other fuel effects studies.

The fuel properties in testing matrix used in the evaluation are in two respects quite different from previous studies. The higher levels of ethanol, up to 20%, are beyond what any previous study has considered. Also, the lack of olefins evaluation data eliminated one fuel parameter that was found to have a significant effect on emissions in previous studies.

One issue to be determined is how EPA or others will choose to extend the fuel property relationships developed in this report to the general fleet, such as in MOVES. Because vehicles will naturally age and may respond differently to fuel properties, these relationships may not continue to hold true. These higher emitting vehicles could potentially contribute to the emissions inventories out of proportion to their numbers. Care should be taken when extending the fuel effects to other vehicles, whether aged late model light-duty or heavy-duty gasoline powered vehicles.

OVERALL REVIEW

The approach to modeling is well considered given the low emissions rates of these vehicles and the relatively small fuel effects, often below normal detection limits. It was apparent that the iterative process using not only novel statistical techniques (compared with other fuel effects evaluations), but an understanding of the testing limitations (low emission rates coupled with detection limits) was a necessary method to determine relevant fuel parameters to include in the evaluation.

I agree with the approach of limiting the number of statistically fit terms (especially the second order terms) to only those that assist in explaining the fuel effects. Perhaps the discussion of the magnitude of the residuals could further highlight the lack of impact of terms that have been dropped from the correlations such as in Figure 70, where a trend may exist, but is relatively small in magnitude.

The overall statistical approach is sound using several methods to discover and systematically eliminate terms. The approach to identify influential data used appropriate physical (investigate detection limits and other laboratory variables) and statistical (significance level, BIC, and residual evaluation) methods to eliminate or keep data.

SPECIFIC COMMENTS

I was unable to find the earlier reports "EPAct/V2/E-89" referenced in the document, presumably "Assessing the Effect of Five Gasoline Properties on Exhaust Emissions from Light-Duty Vehicles certified to Tier-2 Standards: Part I - Study Design and Execution (EPAct/V2/E-89 Test Program Final Report)". This report likely describes in better detail why the fuel blending program could not produce an orthogonal fuel matrix. Likewise, questions about the measurement, vehicle conditioning, and other issues are not addressed in Part II. This makes it difficult to assess the relevance of the modeling.

For example, this statement on Page 36 ("*The alternate measure "NMHC as measured by FID" ($NMHC_{FID}$), was collected for the entire dataset, and it very tightly correlated with both NMOG and "true" NMHC. It is thus possible to estimate NMOG and NMHC results for tests without speciation by using correlations generated from those with speciation. This technique essentially estimates the offset between the response of the FID and the fully characterized emission stream, due to the incomplete measurement of oxygenates by the FID. For NMOG, this estimated value is typically between 2-20% higher than the $NMHC_{FID}$ measurement, depending on emission bag and fuel ethanol level."*) refers to an apparently unique measurement $NMHC_{FID}$ and data handling approach, and it would be useful to understand this measurement in order to understand the validity of this statement. The correlation (described in section 3.2 Imputation of Speciated Hydrocarbons (NMOG, NMHC)) of $NMHC_{FID}$ to NMOG and NMHC appears to indicate that fuel ethanol level has no effect on this correlation (same slope for all levels of ethanol, Tables 10, 11 & 13, 14) for Bag 2 and 3, but an ethanol slope term for Bag 1. Yet, the correlations of acetaldehyde, formaldehyde, and ethanol in Table 82 for Bag 2 demonstrate ethanol still increases these oxygenated species and so should increase NMOG (consisting usually of NMHC + oxygenated carbon). Perhaps there is something unique about this $NMHC_{FID}$ measurement that allows oxygenated species to be measured at some level. However, the approach to correlating NMOG and NMHC with $NMHC_{FID}$ appears inconsistent for Bags 2 and 3 compared to Bag 1.

EDITORIAL SUGGESTIONS

There are numerous editorial corrections that need to be made, but most have already been noted in the document itself or are obvious from WORD program review. What follows are specific suggestions either for improving the flow or for feedback that would not otherwise be considered during normal editing. The original line is provided in *"quoted italics"* and the suggestions, questions, or comments appear in suggested edits (strikethrough or red) or flat text.

Page 9

Statement "*The analysis involved ongoing and iterative interaction between statistical modeling and additional physical and chemical review of the data.*" Page 23 *"The final design is the result of an iterative process involving interactions between research goals, the feasibility of fuel blending, and experimental design."*

Comment The term "interaction" should refer only to statistically fitted second order terms where fuel parameters are mixed, such as ZZ_{ea}, to represent the second order ethanol x aromatics term. The statement above appears to use "interaction" in a different context and so is confusing. EPA should also search all other uses of "interaction" to ensure that there is no confusion.

Statement "*The models reported in this section are as ~~parsimonious~~ concise as the data and subject-matter knowledge allow.*"

Page 13

Statement "*Ethanol: taken in isolation, the models indicate that increasing ethanol is associated with increases in all emissions, both for cold-start and hot-running emissions. The sole exception to the pattern is CO, for which the response to added ethanol is lower CO emissions during start but inconclusive for running conditions. ~~fuel properties appears to change between start and running.~~*"

Statement "*Note that this generalization does not account for the effect of interactions between RVP and other properties, which are in some cases larger than the underlying linear effects.*"

Comment I see only two RVP interacting terms affecting only running THC (not NMOG or NMHC) and start CO, and this statement only appears to be true for the CO start emissions. I would suggest either stating this more plainly or striking the comment. Or does this line refer to fuel properties that are affected when RVP is modified? For example, reduced T50 or diluted aromatics occur when lighter compounds are added to increase RVP; then this statement would be true for the other fuel properties than just for RVP.

Page 22

[5] *"This parameter was measured as DVPE, but for simplicity and consistency, we will refer to it as "RVP.""*

Page 181

Statement *"[14] The typical hydrocarbon analyzer used for emission testing uses a flame ionization detector (FID), which is calibrated to accurately count carbon atoms that are bonded to hydrogen. Carbons bonded to oxygen, which occur in carbonyl and alcohol emissions from burning ethanol fuels, are not accurately counted by the FID, and thus emissions from ethanol fuels require additional characterization methods to properly quantify as NMOG or VOC."*

Comment This statement is generally, but not strictly, accurate in that most FID units are calibrated on propane, and the hydrocarbon measurement assumes that all carbons in the sample respond the same as the carbon atoms in propane. Hydrocarbons do not necessarily respond identically as propane carbon atoms do, but carbons bound to oxygen respond at a rate order(s) of magnitude lower. The error for most hydrocarbons has been considered insignificant. The response to oxygenate carbons has been historically considered to be insignificant compared with other hydrocarbons in the sample, so carbonyl and alcohol compounds determined through alternative methods are added to the NMHC weight measure as the weight of single carbon aldehyde (formaldehyde) and alcohol (methanol). The different composition of hydrocarbons and any FID response to oxygen-bound carbon influence the NMHC measurement, but those influences are considered minor. With lower emission rates of
these newer vehicles, these assumptions may not be as appropriate as for older vehicle designs with higher emission rates. In addition, does the 'typical hydrocarbon analyzer' differ from the $NMHC_{FID}$ measurement that needs to be correlated with NMHC before the statistically modeling proceeds?

Page 182

"8.2.2.4 Vapor Pressure"

Comment I found this section description (and others that rationalize the effect modeled) more of an unsatisfying hand waving exercise to justify what the data was telling. If indeed the older studies and this most recent evaluation are to be believed, then the modeled effect may be temperature dependent when at high temperature the feedback of vapors to the engine may affect the emissions opposite to that modeled here. The suggestion is to limit the speculation to the results of the analysis presented and note other references if the effect found needs to be justified.

Brian West, ORNL, westbh@ornl.gov 865-946-1231 **January 2012**

Review of **"**Assessing the Effect of Five Gasoline Properties on Exhaust Emissions from Light-Duty Vehicles certified to Tier-2 Standards (EPAct/V2/E-89: Phase 3) - Part II: Data Analysis and Model Development - DRAFT REPORT (December 2011)"

The report is long and contains tremendous detail regarding the statistics and modeling approaches used in analysis of the EPAct/V2 data from the SwRI program. Considering its length and the technical detail, the report is well written and well organized. Several recommended edits and comments are provided in the marked up report, to correct typos or improve clarity. As an example, many figures could use larger fonts before final publication. Several cited references (especially Part I of the same report and the appendices) appear unavailable at present, and there are a few missing figures and references. But overall it is a very good draft.

This reviewer has over 20 years experience in engine, emissions, and vehicle testing, but limited experience with many of the statistical methods and modeling approaches described. To the extent possible in the short time available, methods and terms were researched and explored while reviewing the EPA document. Explanations and approaches appear largely reasonable. Some specific questions are noted in comments in the report. For example, the Bayesian Information Criterion (BIC) is used to determine goodness of fit. In one example, the BIC for the full model starts out at >2900 and decreases by 0.1-0.2% for subsequent reduced models. The discussion refers to "a steady decrease" in BIC. The reviewer suggests the significance of such a small change in BIC be discussed further. While the reduction in BIC appears *insignificant*, the fact that BIC is *not increasing* is perhaps the justification to use the reduced model (sec. 5.3.2).

The authors appear to have taken great care to ensure that models are as representative as possible, and to avoid overly complex models or overfitting. The handling of "nondetects" appears reasonable. There is rigorous treatment and discussion surrounding background measurements, analyzer drift, limits of quantitation, censoring of data, etc. Nonetheless, the complexity of the problem makes it difficult to see whether all of the objectives were achieved or rather that some of the apparent results are any more than artifacts of an intricate math problem. As an example, all emissions data are shown by vehicle and include all tests on all fuels, to show the range of measurements for a given pollutant, by vehicle. These are very informative charts. However, no data are shown to demonstrate the test-to-test repeatability for a given vehicle with a given fuel. Several figures (such as Figures 13-21) show averages, but would be more informative if range bars were shown to indicate max and min or perhaps interquartile range, or perhaps scatterplots with all data points.

Authors discuss taking care to not overfit or model random scatter (which is good), in fact stating that the Bag 3 models may be more vulnerable to measurement error due to the extremely low emissions, and opting not to report model coefficients for Bag 3 results. The reviewer agrees with and commends this decision. Handling of extremely low measurements is difficult, and the authors discuss this issue extensively and convincingly. However, when it comes to the PM emissions measurement, there could be additional discussion and justification. For instance, one issue that deserves additional discussion is the lack of tunnel blanks for background PM. Data from a few example tunnel blanks could provide convincing evidence of the zero background assumption. Furthermore, control data should be shown, or at least cited, to support the assumed level of error (± 1 μg) in the PM filter weight measurements. The authors mention discussions with EPA staff as the basis for the assumed level of error. As the authors know, PM measurement is very sensitive to measurement precision and accuracy, with temperature, humidity, buoyancy effects, static charge and other factors greatly influencing results. In SAE 2005-01-0193, the authors detail exhaustive measures to attain an accuracy of better than ± 1 μg in their filter weights. Please provide evidence of the stated PM measurement error.

In Section 7, mean residuals are plotted, presumably to demonstrate the quality of the models. It would be interesting to see range bars on the data points to show the range of individual residuals.

Models should be validated against independent datasets, as discussed in section 8.

Some specific questions were provided for consideration during the review:

1. Was the process of imputation of NMOG/NMHC results for tests/bags with missing speciation data reasonable and statistically sound? (Section X.X) Estimating NMOG from available NMHC appears reasonable and is consistent with what was done in the DOE V4 program (ORNL/TM-2011/234 and ORNL/TM-2011/461). For a given ethanol level, NMOG emissions have been shown to be a linear function of NMHC emissions.

2. Was the decision to remove very low emitting and influential vehicles from the NOx and NMOG analyses reasonable? (Sections 5.5 and 6.1). These discussions were largely reasonable and convincing; however one case (run 6281) was not explained.

3. Please comment on the use of the "design set" of 11 terms as the basis of the final models, versus allowing model to fit all 17 terms, including the adequacy of the justification of this decision. (Section 7). This approach seems reasonable based on the discussion.

4. Comment on the use of the Tobit regression (SAS PROC LIFEREG) for modeling datasets with large numbers of censored values. (Section 5.3). This approach seems reasonable, although I would like to see "large" and "small" censoring levels explained further. Censoring of less than 5 values is considered small. Why? With over 900 datapoints, could the limit be set at a higher number? Please explain.

5. Comment on the decision to independently model the linear terms and interactions between fuel blends as presented in the final results? This approach seems reasonable based on the discussion.

6. Please comment on the methods used to select reduced models. (Section 5). This discussion was convincing, although I am not an expert in this area. I understand that reduced models have lower likelihood of the models describing the random error rather than the underlying fuel effects.

Brian West

Additional comments on the Draft Report

(p.8) The program was conducted in three phases. Phases 1 and 2 were pilot efforts involving measurements on 19 light-duty cars and trucks on three fuels, at two temperatures. This work was completed at Southwest Research Institute between September 2007 and January 2009. ***Have these results been published?***

(p.8) An initial sample of 19 test vehicles was chosen with the intent of representing ~~of~~ the latest-technology light duty vehicles being sold at the time the program was being launched (model year 2008).

(p.9) Speciation also allowed independent analyses of selected toxics including acetaldehyde, formaldehyde, acrolein, benzene and 1,3-butadiene.

(p.9) This approach was followed for several reasons: (1) the candidate fuel effected identified for study were selected because we anticipated that they could be important for one or more emissions. ***Not clear. Candidate fuels? Or candidate fuel effects?***

(p. 10) In generation of emissions, the effects of different fuel properties are not separable, in that it is difficult to modify one property without affecting one or more of the others.

(p. 10) However, the coefficients for different fuel properties can be directly compared, allowing assessment of the relative importance of the effects of the fuel properties on the emissions constituent being modeled.

(p. 13) *Ethanol*: taken in isolation, the models indicate that increasing ethanol is associated with increases in all emissions, both for cold-start and hot-running emissions. The sole exception to the pattern is CO, for which the response to fuel properties appears to change between start and running. The effects are strongest for PM, NOx and NMOG, although presumably, the underlying physical processes could vary. ***Interestingly not consistent with V1 (ORNL/TM-2008/117) or V4 (ORNL/TM-2008/234). Increasing ethanol (in splash blends with certification gasoline) decreased NMHC and THC (FID_HC), while NMOG was relatively flat. Ethanol and acetaldehyde increased, of course. V1 used LA92 but V4 used FTp.***

(p. 13) Cold-start (Bag 1) CO and hot-running THC both have small decreases in emissions with increasing aromatics. In terms of magnitude, the pattern is similar to ethanol, with PM, NOx and NMOG showing the strongest effects.

(p. 16) Since data on Tier 2 vehicles are critical to understanding the impact of fuel property changes on the onroad vehicle fleet as increasing volumes of biofuels are introduced, EPA entered a partnership with DOE and CRC to undertake the largest fuels research program conducted since the Auto/Oil program in the early 1990s. This program is aimed specifically at understanding the effects of fuel property changes on regulated and selected unregulated exhaust emissions from later technology Tier 2 vehicles.

(p. 16) This report describes the analysis of the dataset collected in Phase 3 of the EPAct/V2/E-89 program, conducted at Southwest Research Institute in San Antonio, Texas. A separate report describing the program design and data collection activities is available, but an overview is provided below. ***This report does not appear to be publicly available at this time (January 2012)***

(p. 17) These five parameters were selected based on previous studies on older vehicles as having potential to affect exhaust emissions [cite]. ***Missing reference.***

(p. 17) The parameter ranges to be covered for T50, T90, aromatic content, and RVP were selected to represent the range of in-use fuels based on a review of the Alliance of Automobile Manufacturers' 2006 North American Fuel Survey [cite?]. ***Missing reference.***

(p. 17) Test fuel parameter ranges were originally drafted to span roughly the 5^{th} to 95^{th} percentiles of survey results ~~for~~ in U.S. gasoline, though some test fuel parameters were adjusted after the actual blending process began.

(p. 18) An initial sample of 19 test vehicles was chosen with the intent of representing ~~of~~ the latest-technology light duty vehicles being sold at the time the program was being launched (model year 2008). In terms of regulatory standards, the test sample was to conform on average to Tier 2 Bin 5 exhaust levels and employ a variety of emission control technologies, to be achieved by including a range of vehicle sizes and manufacturers. ***I recall EPA staff indicating (c. 2007) that the list of vehicles was a projection of future technology and/or engine families***

(p. 20) After some consideration, study participants agreed to rely on the aggregate data, while applying appropriate techniques ~~do~~ to address the resulting "censoring" of the data at low end of the range of values. ***Who are the "study participants?"***

(p. 21) The methods used were very similar to those used for the modeling of the other emissions, with modifications to address issues of study design and measurement specific to these compounds.

(p. 22) The design and implementation of the study, including the aspects of fuel blending, measurement methods and logistics are described in a separate report4.**Error! Bookmark not defined.** ***Missing reference.***

(p. 22) It is well known for fuel properties to be moderately to strongly correlated.

(p. 22) [1] This parameter was measured as DVPE, but for simplicity and consistency, we will refer to it as "RVP."

(p. 25) Measurement methods are discussed in detail in the testing report4. ***Missing reference.***

(p. 27) In addition to correlations among the linear effects and interactions, and correlations among interactions, we can see one fairly strong correlation~~s~~ among the linear effects, specifically, between etOH and T50 (R = -0.57).

(p. 31) As one-stage standardization did not neutralize correlations among model terms, we applied a second stage of standardization to the 2nd order terms. ***This report does not appear to be publicly available at this time (Jan 2012)***

(p. 32) Table 8 shows that the combination of one- and two-stage standardization neutralizes the remaining correlations, with the exception that between the etOH and T50 linear effects, as previously described. ***Unclear "that between the etOH and T50 linear effects [what]"***

(p.33) In this constant volume sampling system, the vehicle exhaust is mixed with a large amount of filtered dilution air, and a small portion of this stream is continuously withdrawn to fill a sealed bag over the course of a test cycle.

(p. 33) Its primary disadvantage is that the overall dilution ratio of background air to exhaust must be fixed for an entire test, and is set relatively high to avoid condensation of water vapor within the system during periods of high exhaust flow.

(p. 35) [1] This assessment of the situation ignores the possibility that the vehicle actually consumes or destroys a given pollutant species during parts of the test cycle, resulting in periods of "negative emissions", such that the average emission level over a test is truly zero. While situations may occur over a limited period for some emissions in a highly polluted environment, e.g., PM or NMHC in congested traffic, it is highly unlikely in an emission test cell.

(p. 36) For these reasons, the decision was made not to replace zeros in the dilute bag dataset with integrated continuous measurements. ***Agree with this decision.***

(p. 36) Rather, these measurements ~~to~~ were treated as "censored."

(p. 36) The alternate measure "NMHC as measured by FID" ($NMHC_{FID}$), was collected for the entire dataset, and it very tight~~e~~ly correlated with both NMOG and "true" NMHC.

(pp. 36-37) Based on strong correlations between these species, we developed statistical models to impute NMOG and NMHC from corresponding $NMHC_{FID}$ measurements. ***See ORNL/TM-2011-461***

(p. 37) On this basis, we fit linear models for NMOG and NMHC in terms of $NMHC_{FID}$.

(p. 37) Scatterplots of NMOG vs. $NMHC_{FID}$ for Bag 2 are presented in ***Formatting.***

(p. 38) For the bag 2 and 3 models, the counterpart to Equation 8 is Equation 11, which simplifies for blends other than E0 similarly to Equation 8, except that the slope term is always as in Equation 9.

(p. 50) At the outset, it is helpful to get an overview of the raw results, sorted by vehicle and fuel, which gives an initial impression of variability among vehicles and fuels, as well as within vehicles. ***Agreed. Also would be helpful to show variability of individual vehicles on individual fuels.***

(p. 51) A linear-effects plot for ethanol is shown in Figure 13, which suggests that an ethanol effect is visible when the data ~~is~~ are averaged across the other four fuel properties.

(p. 51) Trends for individual vehicles show a general increase in NOx with increasing ethanol, with some exceptions. ***Consistent with prior studies ORNL/TM-2011/234; SAE 2009-01-2723***

(p. 52) It is necessary to go a step further, and look at "interaction" or "conditional effects" plots, starting with the interaction of ethanol and T50, which deserves special attention because….

(p. 52) In the plot for etOH×T50 (Figure 18), the view seems to indicate an upward trend from E0 through E20, but with some downward curvature above E10. ***Show all data? How much scatter is there in NOx for these cases?***

(p. 52) At first glance, the trend appears to "zig-zag," from low to high. ***Test to test variation of Bag 1 NOx in V1 study (LA92 cycle) ranged from 10% to over 100% (same veh, same fuel).***
In V4 program (FTP, not LA92), range of weighted composite NOx approached 100% in some cases.
How do you ensure that coincidental test-to-test variation is not erroneously attributed to a fuel effect?

(p. 64) As in Bag 1, variability for most individual vehicles spans one third to half an order of magnitude. ***Missing reference?***

(p. 64) The effects for T50 and T90 are similar in that the trends across the five T50 levels and three T90 levels are similar, with no overall trend apparent, for the same reasons noted above for Bag 1 (Figure 26, Figure 27). ***Missing reference?***

(p. 76) The variability within vehicles is about 1-1.25 orders of magnitude. ***Missing reference?***

(p. 76) The Linear Effects plot for ethanol shows some mixed results (Figure 33), but with an apparent increase from 0% to 10% ethanol, followed by a leveling or decline at higher ethanol levels. ***Decline would be expectation. Increase from E0 to E10 is odd***

(p. 77) The Linear Effects plot for T90 is clearly suggestive of an overall positive effect, when considering all vehicles (Figure 37).

(p. 77) The left-hand point represents the same fuel as the left-hand point in the green trend in the previous plot.

(p. 77) In contrast, the plots for ethanol and aromatics are suggestive of a positive or "reinforcement" interaction. In the etOH × arom view (Figure 40), the trend for the higher aromatics level (green) appears steeper than for the lower aromatics level (black). Similarly, in the arom × etOH view (***Font.***

(p. 77) In Figure 41.5, the etOH × T90 plot does not appear to suggest interaction, if we discount the green trend (T90=325°) as representing only two fuels. ***Missing figures***

(p. 91) We assume that a very small but positive measurement existed but was not captured and quantified. Assigning a value of zero to these observations is an example of a common approach to censoring of observations, known as "substitution." In this approach, a small but fixed quantity is substituted for the censored observations. Values used for substitution include zero, as mentioned, or small but positive quantities such as the smallest observation, a multiple of the smallest observation, the limit of quantitation (LOQ) or half the limit of quantitation (LOQ/2). The degree of censoring varied widely by emission and bag, as shown in Table 2. At different stages of the analysis, we addressed censoring in different ways. ***These various approaches all make sense. Suggest adding couple of sentences about the various bags of the LA92. Bag 1 is cold with majority of emissions, hence less censored values. Bag 2 is hot, but includes some open-loop operation due to hard acceleration. Bag 3 is hot start bag with very low emissions, hence majority of censored values.***

(p. 92) Table 25. Numbers of Censored Measurements, by Emission and Bag. ***Suggest showing total number of tests (in title or footnote).***

(p. 93) In this initial step, censored measurements were replaced with the minimum positive value measured for the emission and bag. ***Minimum across all vehicles and fuels?***

(p. 94) As a measure of influence, we calculated the externally studentized or "studentized-deleted" residual (r_{-i}). ***Divided by an estimate***

(p. 94) Table 17. Counts of Influential Measurements, by Emission and Bag (with "influential" defined as having a studentized-deleted residual ≥ 3.5 or ≤ -3.5). ***Suggest including total number of measurements or observations***

(p. 95) An additional measurement in Bag 3 (run 6281) was removed, even though it was not flagged as influential. ***Explain further?***

(p. 95) The full sets of terms in the optimized design include terms anticipated to be meaningful for any of the emissions to be measured. However, it was not anticipated that all the terms included would necessarily be meaningful for all emissions in all bags. A closely related goal is to develop models that would be, to the extent possible, explicable in terms of knowledge of the relevant physical and chemical processes. Parsimonious models are preferred over full models for this purpose, as their simpler structure makes their behavior easier to assess and explain. Finally, with respect to explicability, it is much preferred to minimize the potential for overfitting, which could reduce the generality of models selected for prediction. To guide the process, we adopted several assumptions, described below. ***Model describes the random error rather than the desired underlying relationship. Good discussion.***

(p. 96) For minimal levels of censoring, defined as five or fewer censored measurements ($n_{\text{censored}} \leq 5$), we elected to substitute the minimum positive measured value for the missing measurements. After substitution we fit mixed models as described above. ***N<5 seems reasonable if Total number of samples is large. Suggest noting Ntotal here. Why 5? That is, why not 4 or 6 or 10? Is 5 arbitrary or is there a citable reference to why 5?***

(p. 99) At each step, we tested the goodness-of-fit of each reduced model against that of the full model using a likelihood-ratio test.

(p. 100) In this case, the BIC declines steadily as terms are removed, indicating an improvement in fit for each successive reduced model. ***Declining steadily by 0.1 to 0.2% does not seem significant or important. Is the improvement cited below truly due to the simpler model? Perhaps, but is the BIC really an accurate indicator?***

(p. 101) Table 20. Model Fitting History for PM, Bag 1 (FM9 selected as best-fit model). ***Is the sensitivity of BIC such that there is a significant difference between 2862 and 2867? Six significant figures and***

0.1% change in BIC does not seem important. If this is important or significant, it should be explained. Perhaps the important point is that BIC is NOT INCREASING as the model is simplified, thus justifying the simpler model?

(p. 102) All models based on the 16-parameter full model, as shown in Equation 14 are.....

(p. 107) The approach to analysis of censored measurements, as described in 5.3.2, was also adopted based on guidance from the author of the DOE research. ***No public documents can be found on the EPAct/V2 study.***

(p. 107) Thus, in running these models, censored values were replaced ~~in~~ with the minimum positive measured value in each bag for each emission.

(p. 116) This observation led to a closer examination of measurement error in the dataset.

(p. 116) If this program were simply trying to quantify the magnitude of NOx emissions from such vehicles, this level of error may be acceptable.

(p. 116) However, since we are looking for meaningful differences in emissions between fuels, this large relative error is particularly problematic. ***Key point. Meaningful differences.***

(p. 117) This condition would be expected to give a zero result as discussed in Section 3.1.

(p. 117) This suggests it likely has higher measurement noise than data from the other vehicles, and thus many measurements may not be reliably distinguishable from background levels. ***Good***

(p.118) The data for the Sienna, though similar in their range of sample and background measurements, ~~was~~ were not found to be exceptionally influential to model-fitting and therefore ~~was~~ were not removed from the dataset.

(p. 120) After the test is conducted, the filters are removed, placed back into the clean containers, and returned to the clean room.

(p. 120) In this program, dilution air was HEPA-filtered and presumed to be free of PM, so there was no background filter sample collected for later subtraction as is typical with other emissions. ***Why no tunnel blanks to prove this assumption?***

(p. 120) Discussion with EPA staff experienced with PM measurement suggests that for the data as collected in this program a variability of ±1 µg should be applied to all filter weights. Considering that the net PM result is calculated by subtracting two filter weights (average dirty minus average clean), it should be understood to have a variability range of ±2 µg, as the measurement error applies to both weights. Therefore, a net weight gain of 10 µg would have a relative error of 20% associated with it, a figure of the same order of magnitude as the fuel effects this program attempts to capture. ***Need to cite reference(s) or present data to establish this level of error. For example 2005-01-0193.***

(p. 124) Figure 5~~6~~4 and Figure 57~~5~~ show net measurements as a percentage of sample for bags 2 and 3, as an attempt to understand the magnitude of the nets relative to the measurement error where nets are small.

(p. 124) Two vehicles (Odyssey, Sienna) have the majority of points below those of other vehicles.

(p. 124) Plots of ambient and sample for bags 1-3~~2~~ are shown in Figure 58, Figure 59, and Figure 60, followed by examination of nets as percentage of sample for bags 2 and 3 in Figure 61 and Figure 62.

(p. 138) At this point it is important to note that the results reported below differ from those reported above in 5.3, as well as from those reported for the DOE analysis8. ***Missing reference***

(p. 149) Interaction plots for selected terms are shown in Figure . ***Y axis title is "mean measurement." Isn't this a modeled result, not a measured result?***

(p. 149) It is interesting to note that while (d) and (f) appear somewhat similar visually, the model considers the aromxT90 interaction highly significant but the RVP×T90 interaction insignificant. ***Explain***

(p. 149) Another way of viewing the interactions is to average and plot the residuals of the linear effects model. ***Compare modeled results to measured results. Mean is good, but would also be nice to see the range of residuals. Mean can be close to zero, but how much variation is there? (average of + 20 and -20 is zero).***

(p. 150) < insert physical interpretation of these interactions here?> ***Yes***

(p. 151) Figure 65. ln(CO) (Bag 1): Two-way Conditional Effects Plots for Three Interactions and one Quadratic term, viewed with respect to Both fuel Parameters : (a) Ethanol×Aromatics, (b) Aromatics×Ethanol, (c) Aromatics×T90, (d) T90×Aromatics, (e) RVP×T90, (f) T90×RVP, (g) etOH×etOH. ***Typo in fig c and e. T90 blue should be 300***

(p. 153) Figure 66. CO (Bag 1): Mean Residuals for the Linear Effects Model, vs. Target Fuel Properties for three Interactions: (a) Ethanol × Aromatics, (b) Aromatics × T90, and (c) RVP × T90, and the quadratic term etOH×etOH. ***Blue T90=300***

(p. 158) Another possibility is that the effect of T50 in this presentation is masked by variation in RVP and T90 levels across these fuels.

(p. 161) Figure 69. NOx (Bag 1): Mean Residuals for the Linear Effects Model, vs. Target Fuel Properties for four pairs of terms: (a) Ethanol × Aromatics, (b) Aromatics ethanol, (c) Aromatics × T90, (d) T90×Aromatics, (e) RVP × T90, (f) T90 × RVP, (g) ethanol × T50, (h) T50 × etOH. ***Same comment as above. Mean residuals look good. What is range of individual residuals? Perhaps add error bars to show min and max or perhaps interquartile range of residuals?***

(p. 164) The relations among the ethanol, aromatics and RVP coefficients are similar to Bag 1 and to each other, except that the aromatics coefficient is slightly higher than the ethanol coefficient for NMOG, while ethanol is more important ~~to~~ than aromatics for NMHC.

(p. 164) As in Bag 1, the etOHxarom and etOHxT50 interactions are reinforcements, whereas the etOHxRVP interaction is an interference (etOH positive, RVP negative, interaction negative).

(p. 165) One is that the magnitudes of corresponding coefficients are generally larger for Bag 1 than for Bag 2 emissions, suggesting that the effects of fuel properties are more pronounced for "cold start" than for "hot running" emissions. ***As expected!***

(p. 169) When starting with the 16-term model, the reduced model ~~retains~~ includes RVP and T90 linear terms (both insignificant), plus two interactions not included in the design model: arom×T90 and RVP×T90.

(p. 169) The reasons for these differences are not apparent, but it is clear that the relations among NOx, etOH and T50 are complex ***Or nonexistent?***

(p. 169) It may be appropriate to consider whether the Bag 3 results may be more vulnerable to measurement error attributable to low sample measurements relative to background, given the issues with measurement discussed in 6.1.1 (page 116). ***Yes***

(p. 176) *Ethanol*: taken in isolation, the models indicate that increasing ethanol is associated with increase in all emissions, both in bags 1 and 2.

(p. 178) This question does not arise in the context of model application, but rather with respect to model validation, in that the datasets available for validation include data that represent emissions from pre Tier-2 vehicles. ***Such as the DOE V1 and/or V4 datasets? Have the models been compared/validated against V1 or V4 data? Note that V1 used splash blends and ran the LA92 cycle. V4 tests also used splash blends but ran the FTP.***

(p. 180) The results of the present study are consistent, showing an increase in NOx emissions with an increase in ethanol level, regardless of whether it is from a statistical analysis of an orthogonal change in ethanol alone or when changes in other fuel properties typical of splash or match blending are included.

(pp. 180-81) Thus, despite much lower overall emission levels that have been achieved in Tier 2 vehicles through improved fuel control and catalyst efficiency, the effect of ethanol on combustion and aftertreatment(?) appears to persist in certain modes of operation such as cold-starts and transients during warmed-up operation. ***Effect of ethanol on NOx emissions may be related to exhaust stoichiometry and catalyst efficiency more so than (or in addition to) changes in engine-out NOx.***

(p. 181) [Should we address effect on THC/NMHC as well?] ***Yes. Note that NMOG in V4 was not affected by ethanol, but FID_HC and NMHC decreased (with splash blends).***

(p. 181) In the present study NMOG decreased with decreasing aromatics content, in agreement with earlier studies. ***Or increased with increasing aromatics…***

(p. 182) The present study shows little or no effect of T90 on NOx, which is reasonable considering that NOx production and control are largely about heat release, and T90 represents a smaller amount of combustible material at cold-start temperatures, and thus less energy, than T50. ***Not clear, please review/reword.***

(p. 182) However, during cold start, the new data suggests the direction of the effect is dependent on ethanol content.

(p. 183) Consider showing quantitative parameter changes and percent change results for the models?) ***Good idea***

(p. 184) More variation is seen in T50 and RVP, which span somewhat wider ranges of 160-195 °F and 7.3-13.0 ~~lb~~ psi, respectively.

(p. 184) Use of the non-standardized models is much more straightforward, in that they allow input of fuel properties in their original units, e.g., % for ethanol and aromatics, ~~lb~~ psi for RVP and °F for T50 and T90.

(p. 185) The processes and methods for hydrocarbon speciation are described in greater detail in the testing report4. ***Missing reference***

(p. 187) The "*G*-efficiency" for the full design was estimated at 51.6% for the eleven design parameters, as previously described in 2.1 ~~2.1~~ above (page 22).

(p. 187) As shown in the table, the design efficiency drops sharply with inclusion of 2nd order terms, to less than 5% for designs four and five.

(p. 190) In addition to estimating the sample measurements, as shown above, we estimated two variances, the first being the variance of the ~~variance of the~~ 5-day moving average of the media blanks ($\hat{\sigma}_k^2$), and the second being a variance of random errors ($\hat{\sigma}_\varepsilon^2$).

(p. 193) The first two plots (Figure 8~~12~~) show acetaldehyde vs. ethanol, by T50 level, in linear and logarithmic space.

(p. 199) With respect to censoring, the following rule was applied. If the number of censored measurements was ≤ 5, we substituted the smallest measured positive value for the missing values, and proceeded with model fitting, using a mixed-model approach. ***Why 5? Why not 4 or 6 or 10? Explain.***

(p. 199) However, if the number of censored measurements was > 5, we fit a model using Tobit regression (i.e., "censored normal regression"), an established technique for analysis of left-censored datasets. ***Ditto. What is significance of 5?***

(p. 199) For compounds affected by media contamination, we integrated the approach to censoring with an "Estimated Dependent Variable Model," an approach to modeling datasets with measurement uncertainty in the response variable. ***Ref 20***

(p. 202) If the *p*-value for the test-against-previous is greater than the critical value, the null hypothesis of no significant differenc~~e~~t in fit between the reference and nested models is retained, and the nested model is retained as the current best fit.

(p. 203) Based on these results, the reduced model FM7 was selected as the best fit. ***As noted in previous BIC discussion, do small variations in BIC truly indicate a difference in fit? FM5, FM6, and FM7 all have similar BIC.***

(p. 208) Finally, to illustrate the results of the jackknife replication procedure, Figure 83 shows cumulative distributions of coefficients for each jackknife replicate for the five ~~linear~~ linear eEffects.

(p. 208) The distributions for aromatics (Z_a),T50 (Z_5), and T90 (Z_9) are similar in that they show noticeable lengthening in the lower tail, suggesting that 2-3 vehicles may be influential in decreasing the values of the coefficients.

(p. 211) Detailed results, including models fit, fitting histories, coefficients and tests of effect are presented in Appendices Q.3-W.3 for Bag 1 models, and Q.4 - W.4 for Bag 2 models. ***Not available for review***

(p. 214) 10 References: ***Refs 4 and 8 cannot be found***