APPENDIX AVAILABLE ON THE HEI WEB SITE

Communication 17

Advanced Collaborative Emissions Study (ACES) Phase 3A: Characterization of U.S. 2007-Compliant Diesel Engine and Exposure System Operation

Joe L. Mauderly and Jacob D. McDonald

Appendix A. Advanced Collaborative Emissions Study (ACES) Final Plan for Engine Selection

Prepared by the ACES Engine Selection Process Group

This document was prepared by HEI with input from the HEI ACES Oversight and Advisory Committees and approved by the Committees. It was originally published as Appendix A to the Phase 1 report by Dr. Imad Khalek and colleagues (Khalek IA, Bougher TL, Merritt PM. 2009. Phase 1 of the Advanced Collaborative Emissions Study. CRC Report: ACES Phase 1. Coordinating Research Council, Alpharetta, GA. Available at http://www.crcao.org/publications/emissions/index.html. Accessed February 8, 2012).

Correspondence about Appendix A may be addressed to Dr. Maria Costantini, Health Effects Institute, 101 Federal Street, Suite 500, Boston, MA 02110-1817; e-mail: *mcostantini@healtheffects.org.*

Although this document was produced with partial funding by the United States Department of Energy (DOE), under the terms of Contract/Award Number DE-AC26-05NT42429, and certain motor vehicle and engine manufacturers, the opinions expressed herein are those of the authors and do not necessarily reflect the views of DOE or motor vehicle and engine manufacturers. This document has not been reviewed by public or private party institutions, including those that support the Health Effects Institute; therefore, it may not reflect the views of these parties, and no endorsements by them should be inferred.

Advanced Collaborative Emissions Study (ACES)

Final Plan for Engine Selection

Prepared by the ACES Engine Selection Process Group

1. INTRODUCTION

Overview

The Advanced Collaborative Emission Study (ACES) is a product stewardship initiative undertaken to characterize the emissions and assess the potential cancer and noncancer effects of the exhaust of 2007-compliant heavy-duty diesel engines with advanced technologies and control systems. For this study four manufacturers (Caterpillar, Cummins, Detroit Diesel, and Volvo) have voluntarily provided production-ready engines for additional evaluation of emissions beyond that required by law. All four engines are considered to be representative of a class of engines that are currently being marketed in the US. In the first phase of the project, the exhaust composition of these engines is being characterized to verify that the concentration of regulated emissions are at or below the requirements of EPA's 2007 regulation and to provide data on emissions concentrations of a large number of unregulated compounds, some of which may be of concern because of their potential to contribute to adverse health effects. This phase will be followed by a health effects testing phase during which rats and mice will be exposed to dilutions of the exhaust generated by one of the engines.

One engine will be selected for health effects testing based on the results of the Phase 1 emission characterization according to the process described in this document. In developing this Plan, we have been cognizant that all the engines are expected to have quite low emissions, that there would not be substantial differences among the engines, and that the relatively small number of emission tests for each engine (three) will result in very low statistical power to detect small differences between engines.

This Plan for selecting the engine to be used in the ACES health effects studies includes four parts: (i) methods to compare the emissions; (ii) list of emissions components to be used in comparing engines; (iii) criteria for selection of the engine; and (iv) protocol for data analysis and engine selection. A Draft Initial Plan (dated September 27, 2007) was developed in the summer of 2007 and reviewed and approved by the ACES Steering Committee by conference call on October 4, 2007. Subsequently, some issues arose that led to revision of the Draft Initial Plan, resulting in a Final Initial Plan (dated April 11, 2008). Minor modifications were made during the review of the emissions data to clarify certain parts of the plan, which have resulted in this Final Plan for Engine Selection. These changes are summarized in the next section.

Process for Developing the Plan for Engine Selection

The Engine Selection Process Group was set up to guide the process for selecting the engine for health effects testing in ACES. ESPG members were drawn from the HEI ACES Oversight Committee and the CRC ACES Panel, as recommended in the ACES Project Plan, based on their relevant expertise (in engineering chemistry, toxicology, and statistics). Consideration was also given to representation of different government agencies and industries. Representatives from all four manufacturers who contributed engines to ACES are members of the ESPG. A list of members is provided at the end of this section.

This Final Plan for Engine Selection is based on meetings and discussions that started in April 2007. A "straw person" plan developed by the Engine Manufacturers Association (EMA) was discussed at the meeting of the HEI ACES Oversight Committee and the HEI ACES Advisory Committee on April 17,

2007. The EMA proposal was subsequently sent to all Advisory Committee members with a memo asking for their thoughts about an engine selection process. HEI received many helpful comments, including a revised EMA proposal, and suggestions for alternative approaches from several members of the ACES Advisory Committee. These comments were sent to the members of the ESPG for their June 25, 2007, meeting and discussed at the meeting.

The first draft of the Initial Plan (July 13, 2007) was prepared by HEI staff after the ESPG meeting and sent to the ESPG for comments to assure that it was representative of the breadth of thinking of ESPG members. Based on their comments, including additional explanations of viewpoints, HEI staff revised the first draft. The second draft (August 21, 2007) was sent to the Advisory Committee for its comments. These comments along with input from Oversight Committee members on the ESPG were considered in preparing the Draft Initial Plan (September 27, 2007). That draft was approved by the Steering Committee during a conference call on October 4, 2007. Subsequently comments were received regarding the proposed plan for handling values below the limit of detection (LOD). A small group drawn from members of the HEI ACES Oversight Committee, CRC ACES Panel, and the Phase 1 team at Southwest Research Institute (SwRI) was set up to discuss possible approaches and develop a proposal.

Later, SwRI informed HEI that a decision had been made during the development of the work plan for Phase 1 to eliminate the measurements of dioxins and furans to save costs because of increased costs of other parts of emission testing. This was problematic for two reasons: (1) dioxins and furans were among the emission components that would be used in ranking the engines; and (2) more importantly, these compounds are toxic and their measurement is important whether or not they are used in selecting the engine. The possible measurement of these compounds was discussed with members of the HEI ACES Oversight Committee and CRC ACES Panel. There was general agreement that these measurements should be made. However, because two engines had already been tested when the decision was made, these compounds cannot be used for engine selection. Instead, chloride was added to the list of emission components for engine selection as a surrogate for those species in the Final Initial Plan for Engine Selection. Chloride was selected because furans and dioxins would not be found in emissions unless chlorine was present. Based on a comment from a member of the HEI ACES Oversight Committee, the plan specifies that measurements to be used for engine selection are those made during the 16-hour cycles with crankcase blow-by (see page 4 of this plan). Finally, the list of ESPG members and a section describing the protocol for data analyses and engine selection were added. These changes were incorporated in the Final Initial Plan of April 11, 2008.

This Final Plan contains changes that were proposed by the CRC ACES Panel during the review of the complete emissions data set a meeting in early May, 2008, and subsequently approved by the ESPG. They include: (1) clarifying how to report negative data resulting from background correction; (2) clarifying that total PAHs be referred to as "total PAHs with the exclusion of nitro-and oxygenated PAHs" to avoid double counting of those PAHs, which are in the list of compounds for engine selection on their own; and (3) adding a new rule for removing compounds for which measurements are at or below the LOD in each of the three replicates for three of the four engines. The protocol presented in the April 11, 2008, version has also been updated to expedite the final approval of the engine selection by the Steering Committee. All changes made relative to the September 29, 2007, draft are clearly noted in this Final Plan. Because they did not alter the principles and rules laid out in the Initial Draft, those documents were not sent to the Steering Committee for approval.

Members of the Engine Selection Process Group

From the HEI ACES Oversight Committee: Melvyn Branch (University of Colorado), Kenneth Demerjian (State University of New York, Albany), Helmut Greim (Technical University of Munich), David Kittelson (University of Minnesota), Howard Rockette (University of Pittsburgh), Mark Utell (University of Rochester).

From the CRC ACES Panel (and representatives of each of the 4 companies that contributed engines for ACES): Reynaldo Agama (Caterpillar), Jim Ball (Ford), Steve Cadle (General Motors), Timothy French (Engine Manufactures Association), Tom Hesterberg (Navistar), Donald Keski-Hynnila (Detroit Diesel), Douglas Lawson (National Renewable Energy Laboratory), Hector Maldonado (California Air Resources Board), Mani Natarajan (Marathon Petroleum), Shirish Shimpi (Cummins), Joe Somers (US Environmental Protection Agency), Chris Tennant (Coordinating Research Council), Urban Wass (Volvo), Ken Wright (Conoco-Phillips)

2. FINAL PLAN FOR ENGINE SELECTION

The Plan described in this section presents methods for ranking emission rates of selected emission components from the four engines, emission components to use for comparing emissions, and criteria for selection of one engine for health effects testing, based on ranking of selected engine emission components.

2A. Ranking of Pollutant Emission Rates for Comparison of Engines

The data analyses proposed in this Plan for comparing the engines using the selected emission components (described in section 2B) are based on the following assumptions:

- The study does not have adequate statistical power to distinguish among engines (whose emissions are likely to be similar) because the number of replicate tests needed to detect differences would be outside the budget and timeline of ACES.
- All selected emission components will have the same weight because their relative toxicity can not be assessed at the low concentrations in which they will be present and, therefore, ranking one engine over another based on potential health effects is not possible.
- No engine will have higher or lower emissions for all emission components measured.

Under this scenario traditional statistical techniques that test for differences in individual emission components and adjust for multiple comparisons have limited usefulness. Therefore, we are combining the emission data on selected emission components from each engine into an index that would allow us to rank the four engines and select one for health effects testing.

We present two main approaches for developing an index for ranking emission rates for selected components (see Attachment 1). One is consistent with some nonparametric methods of comparison (straight ranking of measurements) and one is more consistent with a parametric approach (standardized measurements). Each approach has advantages and disadvantages. Examples of the results that would be obtained from each approach are provided in Attachment 1 using an *ad hoc* constructed data set using data from diverse studies and some estimated values to fill in missing data.

I. In the <u>nonparametric approach</u>, each value would be assigned a rank (1, 2, 3, 4, with the sum of the values equal to 10) and the rank values of each emission component for each engine would be added to obtain a total score. An advantage of this approach is that if there are unusually large differences among some values, these tend to be minimized, so it is not sensitive to outliers. At the same time this might be considered a disadvantage if the values are very close, but still assigned different ranks.

II. In the <u>parametric approach,</u> the actual measurements are compared. This approach has the advantage of using all the information available, but because the emission classes have ranges of values that can be considerably different, requires standardization. There are different methods of possible standardization. These include (a) division of each value by the simple mean of the four engines, (b) division of each value by the standard deviation of the mean of the four engines, and (c) subtraction from each value of the mean of the four engines and division by the standard deviation of the mean. With method (c) negative numbers are obtained when the value is lower than the mean of the four values. In all cases the normalized values of each emission component for each engine would be added to obtain a total score.

Method (c) is the typical type of standardization done in statistics, i.e. distance from the average in units of the standard deviation. It is often the preferred method because it uses both the sample mean and the standard deviation and thus provides a reasonable standardization for a wider range of situations. This is also the method that SwRI had proposed for ranking engines.

Examples can be given where the highest (lowest) ranked engine by the straight ranking method is the lowest (highest) ranked using the actual values. However, there are also many samples of data where the two procedures will produce similar ranking. Given the possible variation in ranking with different methods, all four of the methods will be used in the engine selection process (see Section 2C).

Important Issues

1) *How to handle values below the limit of detection (LOD) for emission components for which only some of the measurements are below the LOD*. There are various approaches to handle values below the LOD. A discussion of these approaches and the proposed plan is provided in Attachment 2.

2) *Criterion for defining an engine as an outlier and therefore not representative of other engines.* This is an important but challenging area. As described in Attachment 3, approaches previously considered cannot be used with a very small sample as in ACES (4 engines); the Attachment then goes on to describe a method developed by W.J. Dixon that can detect an engine with extremely low or high measurements. The criterion for an engine to be designated as an outlier is to have at least 50% of the emission components on the selected list be outliers by the Dixon method.

3) *Using tunnel blanks to correct emission measurements.* Regulated emissions will be measured and corrected according to prescribed methods in the Code of Federal Regulations for engine certification, from the EPA, or other methods if approved by the CRC ACES Panel. Unregulated emissions measurements would not be corrected using tunnel blanks, but the values of the tunnel blanks would be reported along with the values of the emission components measured when the engine is running. Because the tunnel was cleaned before the beginning of the ACES testing, and these new engines with control technologies will not coat the tunnel with soot like older engines did, tunnel background levels should be quite low. However, the levels of some of the unregulated species being measured will likely be very low, too, and it is possible that routine subtraction of the tunnel blanks would lead to some negative values. Therefore the final judgment on whether or not any background correction will be applied from the tunnel blanks was left to the CRC ACES panel. This plan clarifies how to deal with negative numbers resulting from background subtraction (see page 13).

2B. Selection of Emission Components for Comparison of Engines

Criteria for Selecting Components for the Initial List

The list of emission components to use for comparing engines includes both regulated and unregulated emissions, including individual compounds, classes of compounds, and physical measurements (particle number). The emissions measurements used will be from 3 replicates of a 16-hour cycle with crankcase blow-by, which will be used for animal exposures and includes both the

FTP and ARB cycles. Where classes of compounds are used in the list of emission components, emission rates of all the compounds in that class will be added together. The criteria for selecting emission components were:

1) Choosing emission components that have toxic properties at some exposure concentration;
2) Avoiding to the extent possible selection of overlapping classes of emission components, which would lead to weighting some categories of emissions more than others (e.g., total PM, hydrocarbons, sulfate, nitrate, and metals would provide 5 measures of PM);
3) Selecting a relatively small number of categories of emission components from the many possible candidates because with a large number of categories none would contribute much.

Initial List of Emission Components (compounds, classes, physical measurements)

- Nitrogen oxides
  - NO, g/hp-hr (1)
  - NO_2, g/hp-hr (2)
    (NO_2 is especially important because of its potential for causing short-term health effects, though levels will be lower in the 2010 engine emissions).
- Carbon monoxide, g/hp-hr (3)
- Particulate matter
  - PM mass, g/hp-hr (4)
  - PM number, particles/hp-hr (5)
  - PM components:
    - Elemental carbon, mg/hp-hr (6)
    - Organic carbon, mg/hp-hr (7)
    - Total metals, μg/hp-hr (8)
    - Inorganic ionic species (ammonium, nitrate, sulfate, chloride), mg/hp-hr (9). (Note: Chloride was added in the Final Initial Plan of April 11, 2008)
- Hydrocarbons
  - Carbonyl compounds, mg/hp-hr (10)
  - Single ring aromatic compounds, μg/hp-hr (11)
  - Polycyclic aromatic hydrocarbons (volatile and particulate)
    - Total PAHs (with the exclusion of nitro and oxygenated PAHs), μg/hp-hr (12) (Note: revised class of compounds in this plan)
    - Total nitro PAHs, μg/hp-hr (13)
    - Total oxygenated PAHs, μg/hp-hr (14)
  - Nitrosamines, μg/hp-hr (15)
  - 1,3-Butadiene, mg/hp-hr (16)
  - Note: dioxin and furans were removed from the list in the Final Initial Plan of April 11, 2008, because these measurements were not available for all four engines.

Criteria for Adding or Deleting Components from the Initial List after Evaluation of Emissions Data

After all emissions data from all four engines are available and reviewed, the initial list of emission components will be evaluated and finalized. Reasons for deleting or adding some emission components are:

1) Emission components for which measurements are at or below the LOD in at least two of each of the three replicates for all four engines would be removed from the list.

2) Emission components for which measurements are at or below the LOD in each of three replicates for three engines would be removed from the list. *If this rule is applied to more than 10% of the total compounds to be used for engine selection, the data will be investigated to understand the possible reasons for this phenomenon before making a decision as to whether the compounds should be removed from the list*. (Note: new rule in this plan.)
3) Individual species would be substituted for classes of compounds for which criterion (1) is met for most species. For example, in the class of carbonyl compounds, formaldehyde may be the only compound detected in any of the samples; if so, formaldehyde may be substituted for carbonyl compounds in the list. This may be the case with other classes of emission components also.
4) If significant, unanticipated results are found upon review of the emission data, emission components may be added to the list upon recommendation of the HEI ACES Oversight Committee in consultation with the ESPG and agreement of the ACES Steering Committee.

Correlations among Emission Components

Information about correlation of various emission components or classes, based on the ongoing engine testing, may provide interesting approaches to be considered in the broader evaluation of emissions results from Phase 1, but would not be used in adjusting the list of emission components. It would be useful to:

- Compare results of the FTP and ARB cycles with the combined 16-hour cycle to see whether there are differences that would not be expected.
- Compare patterns in emissions from different engines to understand better the effects of emission control systems. For example, would an engine with higher PM emissions than other engines have lower levels of NO_2, carbonyl compounds, volatile PAHs, or organic nitrogen compounds and vice versa? Given the expected small variation in emissions across the engines, these comparisons may be informative in identifying trades offs between selected gaseous and particulate constituents.

2C. Method for Engine Selection

Random selection would be a justifiable approach to engine selection if the four ACES engines, as expected, turn out to have very similar emissions. However, in order to maintain the credibility of the study, it is important to be conservative in selecting the engine even though the differences seen in emission rates among the four candidate engines are likely due to random variability, given the small number of measurements. Thus, it is important to avoid choosing an engine with the lowest emissions, which may lead to questions from others about the health effects of emissions of the other engines with higher average emissions. This selection process has been finalized before emission characterization data are available; it will allow for some technical revision in the list of emission components to be used for engine selection, as described in section 2B.

The approach proposed for selecting the engine for health effects testing among the four engines is based on the following assumptions.

1) All engines will meet the 2007 federal emissions standards and both their regulated and unregulated emissions are expected to be similar in concentration and therefore should be considered representative of the engines of the same model.
2) The identity of the selected engine would not be disclosed until the end of the study with publication of the final report in the 2012 timeframe. Also, at that time a new generation of 2010-compliant engines with lower NOx levels in emissions will be on the market.

3) It is unlikely that there will be substantial or systematic differences across the engines.

The engine selection process will use all 4 of the ranking methods described in Section 2B. If 3 or 4 of the methods yield the same 4th ranking engine (i.e., that with the highest emissions index), that engine would be selected for Phase 3 health effects testing. If fewer than 3 methods identify the same 4th ranking engine, then the engine for health effects testing would be selected randomly. The rationale is that this disagreement among methods would provide a factual basis for establishing that the study does not have the statistical power to distinguish among the 4 engines.

Additionally, there are two circumstances in which a particular engine would not be included in the random selection process:

1) If an engine is defined as an outlier (either with high or low emissions) that is not representative of the other engines; or

2) If 3 or 4 of the ranking methods yield the same 1st ranking engine, it should not be included in the random selection because the engine with the lowest emissions (differing significantly from the others) should not be used for health effects studies as indicated earlier in Section 2C.

3. PROTOCOL FOR DATA ANALYSIS AND ENGINE SELECTION

The following protocol is proposed for the purpose of analyzing and presenting the emission data for engine selection only (full analyses of all emissions data will be conducted as part of Phase 1 for CRC). The data analyses for engine selection purposes will be conducted by Drs. Imad Khalek and Robert Mason of SwRI. Engines will be identified as A, B, C, and D.

Data presentation and analysis

1) Emissions data for the compounds and classes of compounds, listed in Section 2B of the Initial Plan for Engine Selection will be tabulated and presented together with their respective LODs for each 16-hour cycle for each engine (Table 1). The units that will be used to report the emission rates are reported in the Initial List of Emission Components on page 5.

2) The measurements will be compared to the LODs and those emissions components with at least 2 of the three replicates for each engine below the LOD or those with all three replicates for 3 engines below the LOD will be struck out. Measurements reported as <LOD will be assigned a value equal to half the LOD (as illustrated in Table 2.1 of Attachment 2). (For measurements below the LOD for which actual values are reported, those values will be used as described in Attachment 2.) For classes of compounds, emission rates for each compound will be added to obtain a final emission rate. Measurements that appear suspect will be flagged and a note will be added to explain. The table containing these data will be referred to as <u>Emissions Data for Engine Selection (Table 2)</u>. (Tables 1 and 2 may be combined.)

3) The mean and standard deviation of the emission rates for the single compounds and for classes of compounds for the three replicate 16-hour cycles for each engine will be determined. This table will be referred to as <u>Summary Emissions Data for Engine Selection (Table 3)</u>.

4) The Summary Emissions Data for all emissions components will be transferred to separate spreadsheets for ranking as described in Attachment 1. The four resulting <u>Ranking Tables for Engine Selection (4-7)</u>, as well as Tables 1, 2, and 3, will be distributed to the ESPG. These analyses will be completed by the SwRI staff within two weeks of receiving all the compositional results.

Engine selection procedure

The following protocol is proposed for making a decision on which engine to select.

5) The tables generated by SwRI staff will be sent to HEI staff who will distribute them to the members of the ESPG.

6) Members of the ESPG will review the data generated by SwRI and participate in a meeting to be held on June 6, 2008 to:

   - Review the data, discuss possible measurement outliers, decide whether other emission components not included in the list should be added, and decide whether any engine meets the criteria for being considered an outlier.
   - Agree on which engine should be selected. If additional analyses were to be recommended, the ESPG would reconvene immediately after these analyses have been conducted to make the final selection.

7. HEI staff will develop a draft report of the decision process and the ESPG recommendation (ESPG Draft Engine Selection Report) within 5 days from the meeting. The report will include Tables 1, 2, 3 and 4. The ESPG will be asked to review the report and provide comments within 5 working days. The HEI staff will finalize the report within 3 days (ESPG Engine Selection Report). This report will be sent on to the ACES HEI Oversight Committee and the ACES Advisory Committee for their information and to the Steering Committee for approval of the process that the ESPG used to implement the Engine Selection Plan. (Note: the steps for review of the ESPG Engine Selection Report have been compressed relative to the version of the protocol in the April 11, 2008, Final Initial Plan in recognition of the substantial agreement of the ESPG about the process and to expedite the process of approval of the engine in order to begin the installation and testing of engine testing at the health effects facility.

8. HEI will organize a conference call or meeting of the Steering Committee as soon as it is reasonably possible to review the recommendations and approve the final decision.

ATTACHMENT 1. Examples of Various Approaches to Rank Engines

Note 1: the compounds and classes of compounds are given as examples and are not intended to be representative of those that will be ultimately be selected for engine comparison.
Note 2: NOx was not used in ranking

Table 1.1. Sample table of emission rates for four engines

Emission Components	**Engine A**	**Engine B**	**Engine C**	**Engine D**
PM (mg/mile)[1]	7.0 ± 2.1	26.4 ± 5.5	24.0 ± 6.5	14.4 ± 4.5*
NOx (ppm)[2]	2.6 ± 0.01	1.7 ± 0.02	2.3 ± 0.01	1.4 ± 0.02
THC (g/mi)[1]	0.094 ± 0.002	0.142 ± 0.009	0.111 ± 0.004	0.120 ± 0.005*
PM number[2]	83700 ± 2010	74400 ± 2490	51300 ± 952	655000 ± 2170
formaldehyde (mg/km)[3]	0.12 ± 0.14*	0.06 ± 0.10*	0.09 ± 0.08*	0.11 ± 0.09*
NO (ppm)[2]	0.76 ± 0.006	0.5 ± 0.007	0.72 ± 0.005	0.56 ± 0.007
NO_2 (ppm) [2]	1.80 ± 0.009	1.19 ± 0.013	1.58 ± 0.009	0.84 ± 0.009
total PAH µg/km (PM phase)[3]	4.4 ± 0.7*	4.9 ± 0.5*	3.5 ± 1.0*	3.0 ± 0.7*
Benzo(a)pyrene µg/km (PM phase)[3]	0.07 ± 0.1*	0.22 ± 0.1*	0.17 ± 0.09*	0.47 ± 0.12*
total PAH (semivolatile)[3]	11 ± 3.0*	26 ± 4.8*	35 ± 5.6*	30 ± 4.3*

1. Durbin et al 2003 (Atmos Env)
2. Kittelson et al 2005 (Aerosol Science)
3. Modified from Westerholm and Egeback 1994 (EHP)

* not from papers

Table 1.2. Ranking method 1 (Lowest emission = 1, highest emission = 4, total value for four engines = 10)

Emission components	Engine A		Engine B		Engine C		Engine D	
PM (mg/mile)[1]	7.0 ± 2.1	1	26.4 ± 5.5	4	24.0 ± 6.5	3	14.4 ± 4.5*	2
NOx (ppm)[2]	2.6 ± 0.01		1.7 ± 0.02		2.3 ± 0.01		1.4 ± 0.02	
THC (g/mi)[1]	0.094 ± 0.002	1	0.142 ± 0.009	4	0.111 ± 0.004	2	0.120 ± 0.005*	3
PM number[2]	83700 ± 2010	3	74400 ± 2490	2	51300 ± 952	1	655000 ± 2170	4
formaldehyde (mg/km)[3]	0.12 ± 0.14*	4	0.06 ± 0.10*	1	0.09 ± 0.08*	2	0.11 ± 0.09*	3
NO (ppm)[2]	0.76 ± 0.006	4	0.5 ± 0.007	1	0.72 ± 0.005	3	0.56 ± 0.007	2
NO_2 (ppm)[2]	1.80 ± 0.009	4	1.19 ± 0.013	2	1.58 ± 0.009	3	0.84 ± 0.009	1
total PAH µg/km (PM phase)[3]	4.4 ± 0.7*	3	4.9 ± 0.5*	4	3.5 ± 1.0*	2	3.0 ± 0.7*	1
Benzo(a)pyrene µg/km (PM phase)[3]	0.07 ± 0.1*	1	0.22 ± 0.1*	3	0.17 ± 0.09*	2	0.47 ± 0.12*	4
total PAH (semivolatile)[3]	11 ± 3.0*	1	26 ± 4.8*	2	35 ± 5.6*	4	30 ± 4.3*	3
total ranking		22		23		22		23

Table 1.3. Ranking method 2a (mean of each engine/mean of four engines)

	Engine A	SD		Engine B	SD		Engine C	SD		Engine D	SD		Mean of 4 engines	SD
PM (mg/mile)[1]	7.00	2.10	0.39	26.40	5.50	1.47	24.00	6.50	1.34	14.40	4.50	0.80	17.95	8.95
NOx (ppm)[2]	2.60	0.01		1.70	0.02		2.30	0.01		1.40	0.02		2.00	0.55
THC (g/mi)[1]	0.09	0.00	0.81	0.14	0.01	1.22	0.11	0.00	0.95	0.12	0.01	1.03	0.12	0.02
PM number[2]	83700	2010	0.39	74400	2940	0.34	51300	952	0.24	655000	2170	3.03	216100	292917
formaldehyde (mg/km)[3]	0.12	0.14	1.26	0.06	0.00	0.63	0.09	0.08	0.95	0.11	0.09	1.16	0.10	0.03
NO (ppm)[2]	0.76	0.01	1.20	0.50	0.01	0.79	0.72	0.01	1.13	0.56	0.07	0.88	0.64	0.12
NO2 (ppm)[2]	1.80	0.01	1.33	1.19	0.01	0.88	1.58	0.01	1.17	0.84	0.01	0.62	1.35	0.42
total PAH µg/km (PM phase)[3]	4.40	0.70	1.11	4.90	0.50	1.24	3.50	1.00	0.89	3.00	0.70	0.76	3.95	0.86
Benzo(a)pyrene µg/km (PM phase)[3]	0.07	0.10	0.30	0.22	0.10	0.95	0.17	0.09	0.73	0.47	0.12	2.02	0.23	0.17
total PAH (semivolatile)[3]	11.00	3.00	0.43	26.00	4.80	1.02	35.00	5.60	1.37	30.00	4.30	1.18	25.50	10.34
total ranking			7.22			8.54			8.76			11.48		

Table 1.4. Ranking method 2b (mean of each engine/SD of four engines)

	Engine A	SD		Engine B	SD		Engine C	SD		Engine D	SD		Mean of 4 engines	SD
PM (mg/mile)[1]	7.00	2.10	0.78	26.40	5.50	2.95	24.00	6.50	2.68	14.40	4.50	1.61	17.95	8.95
NOx (ppm)[2]	2.60	0.01		1.70	0.02		2.30	0.01		1.40	0.02		2.00	0.55
THC (g/mi)[1]	0.09	0.00	4.70	0.14	0.01	7.10	0.11	0.00	5.55	0.12	0.01	6.00	0.12	0.02
PM number[2]	83700	2010	0.29	74400	2940	0.25	51300	952	0.18	655000	2170	2.24	216100	292917
formaldehyde (mg/km)[3]	0.12	0.14	4.00	0.06	0.00	2.00	0.09	0.08	3.00	0.11	0.09	3.67	0.10	0.03
NO (ppm)[2]	0.76	0.01	6.61	0.50	0.01	4.35	0.72	0.01	6.26	0.56	0.07	4.87	0.64	0.12
NO2 (ppm)[2]	1.80	0.01	4.25	1.19	0.01	2.81	1.58	0.01	3.73	0.84	0.01	1.98	1.35	0.42
total PAH µg/km (PM phase)[3]	4.40	0.70	5.12	4.90	0.50	5.70	3.50	1.00	4.07	3.00	0.70	3.49	3.95	0.86
Benzo(a)pyrene µg/km (PM phase)[3]	0.07	0.10	0.41	0.22	0.10	1.29	0.17	0.09	1.00	0.47	0.12	2.76	0.23	0.17
total PAH (semivolatile)[3]	11.00	3.00	1.06	26.00	4.80	2.51	35.00	5.60	3.38	30.00	4.30	2.90	25.50	10.34
total ranking			27.21			28.96			29.85			29.52		

Table 1.5. Ranking Method 2c [(mean of each engine minus mean of four engines)/sd of four engines)]

	Engine A	SD		Engine B	SD		Engine C	SD		Engine D	SD		Mean of 4 engines	SD
PM (mg/mile)[1]	7.00	2.10	-1.22	26.40	5.50	0.94	24.00	6.50	0.68	14.40	4.50	-0.40	17.95	8.95
NOx (ppm)[2]	2.60	0.01		1.70	0.02		2.30	0.01		1.40	0.02		2.00	0.55
THC (g/mi)[1]	0.09	0.00	-1.14	0.14	0.01	1.26	0.11	0.00	-0.29	0.12	0.01	0.16	0.12	0.02
PM number[2]	83700	2010	-0.45	74400	2940	-0.48	51300	952	-0.56	655000	2170	1.50	216100	292917
formaldehyde (mg/km)[3]	0.12	0.14	0.83	0.06	0.00	-1.17	0.09	0.08	-0.17	0.11	0.09	0.50	0.10	0.03
NO (ppm)[2]	0.76	0.01	1.09	0.50	0.01	-1.17	0.72	0.01	0.74	0.56	0.07	-0.65	0.64	0.12
NO2 (ppm)[2]	1.80	0.01	1.06	1.19	0.01	-0.38	1.58	0.01	0.54	0.84	0.01	-1.21	1.35	0.42
total PAH µg/km (PM phase)[3]	4.40	0.70	0.52	4.90	0.50	1.10	3.50	1.00	-0.52	3.00	0.70	-1.10	3.95	0.86
Benzo(a)pyrene µg/km (PM phase)[3]	0.07	0.10	-0.96	0.22	0.10	-0.07	0.17	0.09	-0.37	0.47	0.12	1.40	0.23	0.17
total PAH (semivolatile)[3]	11.00	3.00	-1.40	26.00	4.80	0.05	35.00	5.60	0.92	30.00	4.30	0.44	25.50	10.34
total ranking			-1.67			0.08			0.96			0.63		

ATTACHMENT 2. Guidelines for Dealing with Measurements Below the LOD

Introduction

It is likely that the measurements for the 16-hour cycle for some exhaust components to be used for engine selection will have values below the LOD. Methods for addressing these values are complicated due to varying definitions of the LOD, different approaches of treating such values in the computation of a summary statistic, and the fact that SwRI does not report values below the LOD while Desert Research Institute reports the actual values regardless of whether they are above or below the LOD (see below).

Various approaches have been proposed by ACES stakeholders to deal with measurements below the LOD. These include:

- Assigning an arbitrary value (for example LOD/2 or zero). Statistically, this method has the undesirable property that it may result in bias when estimating the mean of several replicates. Similarly the exclusion of values below the LOD results in bias.
- Assigning the value reported by the instrument (either positive or negative). Because measurements below the LOD are considered by some to be unreliable, this method also has limitations. The appropriateness of this method may also be highly dependant on the definition of the LOD.
- Assigning a value estimated by obtaining maximum likelihood estimates. This method is superior to many other methods, but may not be applicable to the ACES data sets because of the small sample size (3 replicates for each compound).
- Assigning a value using a nonparametric ranking approach. Because the engine selection plan relies on both parametric and nonparametric ranking, this method cannot be used.

Each approach may introduce a bias in the estimate of the average exposure, but the engine selection process would not necessarily be biased (i.e. there would be no preference for selecting one engine over another).

Calculation of LOD

SwRI calculates the LOD using different approaches or a combination of approaches. Measurements below the LOD are reported as <LOD. The DRI investigators calculate the LOD as three times the standard deviation of eight replicates of a low-level injection standard and reports the uncertainty associated with the value. Values below the LOD are reported as measured. Zero is used to indicate:

- A true zero reported by the instrument.

  A number which is obtained after data processing, i.e. after subtraction of background or media blanks. If the result is negative after background or blank subtraction, it will be reported as zero. (Note: new rule.)

- For nitro-PAH DRI has to "zero-out" some numbers based on their MS/MS analysis, i.e. when it is determined that a given peak is not a nitro-PAH, but some co-eluting compound, based on the specific fragmentation pattern.
- An invalid number. (From DRI: An invalid number is when something goes wrong during the analysis of the data or the number does not make sense to the laboratory supervisor or analyst. These are usually highlighted in yellow in DRI practices.)

Table 1 indicates, for each compound and class of compounds to be used in engine selection, the laboratory that is conducting the chemical analysis.

Proposed Approach

Recognizing the lack of an optimal method of addressing values below the LOD for our situation we propose the following plan. Examples of the calculations for the different conditions are provided in Table 2.1.

1. If at least two of the three observations for an emission component are below the LOD for all four engines, this component will be dropped from the set of components used for engine selection.
2. If three observations for an emission component are at or below the LOD for three engines, this component would be removed from the list. *If this rule is applied to more than 10% of the total compounds to be used for engine selection, the data will be investigated to understand the possible reasons for this phenomenon before making a decision as to whether the compounds should be removed from the list*. (Note: new rule in this plan.)
3. Individual components:
    a. For measurements below the LOD for which the actual values are reported, those values (positive, zero, and negative) will be used in the computation of the mean and standard deviation.
    b. For measurements below the LOD for which the actual values are not reported, the LOD/2 will be used in the computation of the mean and standard deviation.
4. Classes of compounds:
    a. For measurements below the LOD for which the actual values are reported, those values (positive, zero, and negative) will be used in the computation of the total of that class for each 16-hour cycle. The three totals will be used in the computation of the mean and standard deviation for that class (see example for aromatics in Table 2.1).
    b. For measurements below the LOD for which the actual values are not reported, the LOD/2 will be used in the computation of the total of that class for each 16-hour cycle as described (see example elements in Table 2.1).

Overall Considerations

Although many different algorithms could be used to select an engine (including random selection), the task is viewed as a method of selecting one engine from a group of “equals”. Our emphasis has been to develop the rule a priori and to have an algorithm that is not biased in regard to which engine is selected. Therefore, the procedure for addressing values below the LOD is viewed as a mechanism of facilitating our a priori developed selection procedures and not as applying a method of statistical inference. The proposed method of addressing measurements below the LOD is being suggested for engine selection only.

Table 2.1. Example of Measurements for Engine Selection						
Compound (analytic lab)	How <LOD is reported	LOD	Engine 1			Mean ± SD
NO2 (SwRI)	<LOD					
NO (SwRI)	<LOD					
CO (SwRI)	<LOD					
PM mass (SwRI)	<LOD					
EC (DRI)	actual measure	1.00[A]	1.50	1.80	0.90	1.40 ± 0.46
OC	actual measure					
Elements (SwRI) (class)	<LOD					
Element 1*		1.00[B]	<LOD (0.50)	<LOD (0.50)	<LOD (0.50)	
Element 2		2.00[B]	<LOD (1.00)	2.40	2.30	
Element 3*		4.00[B]	4.10	<LOD (2.00)	<LOD (2.00)	
Element 4		1.00	1.40	1.60	1.30	
Element 5		0.50	2.50	2.70	2.30	
total			9.50	9.20	8.40	9.03 ± 0.57
Carbonyls (SwRI) (class)	<LOD					
Sulfate (SwRI)	<LOD	2.00[C]	<LOD (1.00)	2.40	2.10	1.83 ± 0.74
Nitrate (SwRI)	<LOD					
Ammonia (SwRI)	<LOD					
Single ring aromatics (DF (class)	actual measure					
Species 1		0.50	0.00	0.50	0.60	
Species 2*		0.50[D]	0.20	0.10	0.00	
Species 3		0.70[D]	0.50	0.70	0.80	
Species 4*		0.70[D]	0.30	0.20	0.10	
total*		2.40[D]	1.00	1.50	1.50	1.33 ± 0.29
Total PAH DRI) (class)	actual measure					
Total NPAH (DRI) (class)	actual measure					
Total oxyPAH (DRI) (class)	actual measure					
Nitrosamines (9DRI)	actual measure					

* compound (or class of compounds) has not been removed because a comparison with other engines was not made

A: rule 2.a; B: rule 3.b; C: rule 2.b; D: rule 3.a, see page 12

ATTACHMENT 3. Identification of an Engine as an Outlier

There are many statistical as well as practical issues associated with the identification of "outliers". There are also additional problems that may occur when the sample is small. For example, a commonly used technique is to use the statistic $(x_{(i)} - \bar{x})/s$ where $x_{(i)}$ is a specific observation in a sample of size *n*, $\bar{x}$ is the sample mean and *s* is the sample standard derivation. The x_i value is considered an outlier if the statistic is greater than 3 (corresponding to an observation being a distance of more than three standard derivations from the sample mean). This approach was recommended by some members of the ESPG. Because it can be shown that the value of this statistic cannot exceed $(n-1)/\sqrt{n}$, in our case, with *n*=4, this method would never identify an outlier. The maximum observable difference would be 1.5 *s* from the mean as shown in the following examples.

Example 1
Values: 0.06, 0.09, 0.12, 0.45; mean 0.180 ± 0.182
Highest value (0.45) is 1.48 *s* from the mean

Example 2
Values: 0.06, 0.09, 0.12, 0.95; mean 0.305 ± 0.431
Highest value (0.95) is 1.50 *s* from the mean

Example 3
Values: 0.06, 0.09, 0.12, 2.3; mean 0.642 ± 1.105
Highest value (2.3) is 1.50 *s* from the mean

Given this limitation, one statistician in the ESPG has suggested using a method proposed by Dixon (1,2) to identify an outlier for a particular compound or class of compounds. Let $x_{(i)}$ be the measurements for the four engines ordered from lowest to highest (e.g. $x_{(1)}$ is the minimum measurement, $x_{(4)}$ is the maximum measurement). Then we propose to use the statistic

$$T = \max\left[\frac{x_{(4)} - x_{(3)}}{x_{(4)} - x_{(1)}}, \frac{x_{(2)} - x_{(1)}}{x_{(4)} - x_{(1)}}\right]$$

If $T \geq 0.922$ we would reject at $\alpha = .01$. In applying the technique, we note the following:

1. The method will detect an engine with an extremely low measurement as well as one with an extremely high measurement.
2. If the exposure measurements for a particular compound are known to be more similar to a lognormal than a normal distribution, then the log of the observations should be taken before applying the test.
3. The procedure identifies the outlier for a particular compound or class of compounds. For an engine to be considered an outlier we could use a criterion such as requiring ≥ 50% of the compounds or classes to be identified as an outlier for that engine.
4. A modification to the test may be required if there are multiple values at the LOD.

Example 1 T = 0.85 (highest value is not an outlier)

Example 2 T = 0.93 (highest value is an outlier)

Example 3 T = 0.97 (highest value is an outlier)
(Numbers used in these examples are the same as those used above.)

References

1. Dixon WJ. “Analysis of Extreme Values”, Annuals Math Stat, 21:488-506, 1950.
2. Dixon WJ. “Ratios Involving Extreme Values”, Annals Math Stat. 22:68-78, 1951.