Allen L. Robinson
Fuel Toxics, TOG &PM Speciation Peer Review Comments

General comments

Overall I think that EPA has done a good job of developing MOVES2014 and that these chapters provide the reader/user a reasonable description of the model. The models are statistical fits of data; that is probably the best approach given the limitations in our quantitative understanding in the underlying physical and chemical processes that control the emissions. For the most part the models seem to be based on the best available datasets, but there are inevitably gaps. In certain instance there appear to be important data that are not incorporated into the analysis. I have provided many comments on individual chapters. The majority of the comments are focused on improving the usability of the materials. However, there are some important scientific shortcomings (treatment of uncertainty, semivolatile PM, and SOA precursors).

Here are the major comments that apply across most if not all of the sections that I read:

Presentation related:

1. Data sources – the various chapters and report often provide references to the underlying data. However, these references often point to large reports (e.g. the EPAct data analysis), which means that the reader may not be able to figure out what specific data were used. I would encourage EPA to be as specific as possible about what data are used. I have often been frustrated trying to figure out the exact data underlying models like MOVES and MOBILE.
2. Examples – I think quantitative examples really help the reader understand the model. These exist in a few chapters but not in most. I would encourage EPA to include more examples which will help the reader understand what MOVES is doing. Pointing the reader to online tools, such as the fuel effects spreadsheet are also useful.
3. Tables defining all variables – in some chapter many variables are not defined making it difficult for the reader to understand the model. These tables should also indicate which variables are user inputs and which are derived from existing data. For the user inputs, default values should be clearly defined.
4. Example results – For the reader it would be useful to provide some sample output from the model to understand the effects. Ideally this would be graphical presentation.

Content related:

5. Goodness of fit – Given that the models are statistical fits of data, some description is needed in each chapter on how well the model(s) fit the underlying data is important. These could be some sort of statistical measure and/or scatter plots of model predictions versus underlying data.

6. Uncertainty – There is no discussion of uncertainty of the model predictions. This is my largest substantive concern with the reports. One measure of the uncertainty is the quality of the statistical fit. A better measure is how well the model performs against data that were not used to derive the fitting parameters. I strongly encourage EPA to quantify the uncertainty in the MOVES2014 predictions. Every prediction should be accompanied by a quantitative uncertainty estimate.

7. Data limitations – EPA has done a good job utilizing existing data. However, there are inevitably gaps. Obvious gaps are GDI, higher mileage vehicles, high emitters, etc. The reader should be made aware of these limitations and guidance should be given about how to address.

Review of

TOG and PM Speciation
PM Speciation Appendix
LightDutyEmissionRatesReport_Draft
LightDutyEmissionRatesReport_Draft
HD Emission Rates Report Draft

1. Does the presentation give a description of selected data sources sufficient to allow the reader to form a general view of the quantity, quality and representativeness of data used in the development of emission rates? Are you able to recommend alternate data sources might better allow the model to estimate national or regional default values?

   The report provides some description of data sources. For example Table 12 points the reader to different EPA reports. That is valuable, but it is not clear that the information in the Table is sufficient if a reader wanted to truly understand where the source profile came from. I have been frustrated in the past trying to track down the source data for speciation profiles used in EPA models. Sometimes there are no references (not a problem here), but other times the references point to a large report (the case here). However, these reports can be massive documents that describe lots of data, but the reader has no idea which specific data were actually used to develop the input for the model (or how they were used). Maybe that is not an issue here (I have not gone and looked at the underlying reports), but I would encourage the authors to make sure the reader truly can figure out where the source profiles came from so that can start with the actual data and recreate the actual profiles. For example, the report could refer to specific emissions data form the underlying report.

   The report seems to do a better on the PM side of things (PM speciation appendix, which is built upon this unpublished paper). It is very helpful that the PM appendix includes the actual profiles. I would encourage EPA to write a similar Appendix for the TOG speciation.

   I was surprised that there modeling assumes that a constant EC/PM emission ratio for LDGV. This may be because the KCVES did not test many Tier 2/LEV2 vehicles. The CRC A74/E96 project found a pretty significant increase in the EC/PM for newer Tier 2/LEV2 vehicles. This has been presented in project reports and will be published shortly.

   It also seems like default LDGV EC/PM ratio is not appropriate for GDI, which are becoming a larger part of the fleet. ARB has been doing a fair bit of testing of GDI – presumably those data are available. This will be critical for MOVES to be able to predict emissions from future fleets.

2. Is the description of analytic methods and procedures clear and detailed enough to allow the reader to develop an adequate understanding of the steps taken and assumptions made by EPA

to develop the model inputs? Are examples selected for tables and figures well-chosen and designed to assist the reader in understanding approaches and methods?

3. Are the methods and procedures employed technically appropriate and reasonable, with respect to the relevant disciplines, including physics, chemistry, engineering, mathematics and statistics? Are you able to suggest or recommend alternate approaches that might better achieve the goal of developing accurate and representative model inputs? In making recommendations please distinguish between cases involving reasonable disagreement in adoption of methods as opposed to cases where you conclude that current methods involve specific technical errors.

I like the approach of defining nonECPM because EC is refractory while other components, in particular OC, are semivolatile. This addition is an important step towards implementing a more physically realistic treatment of OC. However, I am concerned that the model continues to treat OC as an inert, non-volatile component of the exhaust. Presumably MOVES is supposed to estimate the PM emissions at typical atmospheric conditions (not those in CVS). The problem is that the low levels of dilution commonly often used in vehicle testing campaigns such as the KCVES create high PM concentrations in the CVS. This bias the gas-particle partitioning of the OC. Few studies have quantified the behavior, but the recent CRC A74/E96 project demonstrates the issues with fleet of 60+ LDGV and MDDV/HDDV vehicles (see May et al. dx.doi.org/10.1021/es400782j | Environ. Sci. Technol. 2013, 47, 8288−8296, May et al. Atmospheric Environment 77 (2013) 128e139). At a minimum the report should point out this limitation that the emission rates may be overestimated because of partitioning biases. I would encourage EPA to start explicating accounting for these biases in both the MOVES emission rates and source profiles. This can be done using the volatility distributions in the May et al. papers and the measured CVS concentrations.

I was confused with section 3 which describes the method for converting between different classes of gas phase organics (NMOG, TOG, THC, etc.).

- First, Title of section 3. Hydrocarbon speciation. I found this confusing. Hydrocarbons are organic compounds that contain carbon and hydrogen. This is a subset of the organic, which can contain compounds in addition to C and H. This should be called total organic gas speciation.
- Second I am concerned with defining the THC emissions based on what is measured by the FID. I realize that this is standard definition but it is not scientifically correct. The FID measures carbon. A problem is that the measurement efficiency is species dependent (as mentioned in the document). The FID quantitatively measure carbons in hydrocarbons (organic compounds comprised of carbon and hydrogen) and the standard propane calibration works well. However, the FID can also measure some of the carbon in oxygenated organics (especially carbons not associated with oxygen atoms) so some of the signal in the FID comes from oxygenated organics, which are not hydrocarbons. Therefore, there is no straightforward interpretation of the FID signal, but it does detect more than just the hydrocarbon emissions.

- Third, I could not follow the equations used to convert between the different classes of organic gases (NMOG to NMHC, etc. – e.g. section 3.2). This correction seems to be relatively straightforward – it appears that you are simply using different ratios of, e.g. NMOG to FID defined THC. Not surprisingly, these ratios depend on vehicle MY and type of fuel.
- I will focus my comments on section 3.2 but the same comments to apply to the other sections (e.g. 3.3) that perform the same analysis. What is the basis of equation (1)? Some underlying physical or chemistry principle? How is equation (1) used? Is equation (1) used to derive un-numbered equations later on page 9? What is the definition CF is molar or mass carbon fraction? MPC is mass of what? per carbon? Where is FIDx defined – give table or reference? Is the speciation constant listed in Table 5 the same as the speciation factor defined by equation 1? If so then you need to reconcile the names. I tried played with equation with equation (1) but could not figure out some of the inputs. It should be clear that I found this whole section pretty confusing and do not have a basic understanding of what MOVES is doing, never mind being able to reproduce the calculations.
- It seems that the key to calculating the needed ratios is not equation (1) but the un-numbered equations listed on page 9. The inputs for these equations appear to be given in Table 4 and 5. I assume that these values are fixed (or can the user input a difference volume to weight percent oxygen)? Where did these values come from? Derived from fuel analyses? Derived from fitting experimental data? If they are fixed, then it seems like one could get rid of Table 4 and simply replace Table 5 with the actual ratios used to convert between NMHC and NMOG for the different model year groups. That would be much simpler. I think that the equations make it appear that what is being done is more sophisticated then it is.

Page 25 "Step 2" states that sulfate and particular water emissions were obtained by speciation profiles. However, I thought these were calculated with the sulfate model?

The report should define what is meant by the ratios of means (or mass weighted means) used to create average profiles. Right now the report assumes the reader can knows this.

4. In areas where EPA has concluded that applicable data is meager or unavailable, and consequently has made assumptions to frame approaches and arrive at solutions, do you agree that the assumptions made are appropriate and reasonable? If not, and you are so able, please suggest alternative sets of assumptions that might lead to more reasonable or accurate model inputs while allowing a reasonable margin of environmental protection.

Limited data for GDI. This is not mentioned in report. ARB has been doing some work on this.

Limited data for CNG. This is acknowledged in the report. Not clear how critical a gap that is given the limited number of CNG vehicles (maybe important in places like LA or NYC with lots of CNG buses?).

Limited data for post-2007 diesels, especially on long-term performance on aftertreatment devices.

These limitations are expensive to address. They should be pointed out in the report.

5. Are the resulting model inputs appropriate, and to the best of your knowledge and experience, reasonably consistent with physical and chemical processes involved in exhaust emissions formation and control? Are the resulting model inputs empirically consistent with the body of data and literature that has come to your attention?

   The PM profiles were weighted using Kansas City MSA VMT data. How sensitive are the profiles to that assumption? If they are sensitive then that potentially creates a number of concerns. How representative is that of other areas in the country? How representative are they of future vehicle fleets?

   Section 4.2 – "But they are the major species by mass and reactivity" I am concerned about the gaps between speciated and total emissions. The standard approach (adopted here), assumes that the unspeciated portion of the NMOG behaves the same as the speciated. This likely is not the case when it comes to secondary organic aerosol (SOA) formation. The unspeciated emissions are likely a complex mixture of higher molecular weight species – these species contribute disproportionately to SOA formation relative to lighter species (e.g. propane).

   "while assuring that the PM2.5 species achieved a 100% mass balance" I find these sorts of statements very concerning, especially given that these sorts of renormalizations are often poorly documented resulting in users not being aware of these assumptions. It is important to document if there are significant mass balance discrepancies, not just normalize them away. I realize that the profiles don't have a PM_unkown species, but enforcing mass balance may create other problems.

Other studies with diesel (e.g. Schauer et al. 1999 EST, Subramanian et al. 2009 EST) show a pretty significant gap in PM mass balance for diesels (sum of speciated low).

Other

Page 5 Intermediate PM section -- EC is not a "real" species in that it is not a distinct chemical substance but something that is operational defined. Although not defined, I assumed a real species was an actual chemical species like CO.

Page 7 Real speciation profile – A key shortcoming is that these real profiles are incomplete – they are typically missing around a quarter of the TOG mass. This point is mentioned later but should be mentioned here as well.

Page 7 – I am concerned with defining the THC emissions based on what is measured by the FID. I realize that this is standard definition but it is not scientifically correct. The FID measures carbon. A problem is that the measurement efficiency is species dependent (as discussed in the document). For FID quantitatively measure carbons in hydrocarbons (organic compounds comprised of carbon and hydrogen) and the standard propane calibration works well. However, the FID will also measure some oxygenated organics (especially carbons not associated with oxygen atoms) so some of the signal in the FID comes from oxygenated organics, which are not hydrocarbons.

The qualifier "start" is often used to characterize the emissions. Every instance of that should be further classified as cold or hot start, as that can make a big difference on emissions. Many times it was not clear what type of start the text was referring too.

Typos

Page 3 defined **by** discrete – missing by

Page 3 although "county"? Not sure what county is

Page 9 "as the all" delete the

Page 14 – "3.1 NMHC and VOC calculations …" this section heading is misnumbered.

PM fractions of median profile greater than 1 → how much greater than 1?

Review of Fuel Sulfur Effects – only Tier 2 sulfur model

General comment – I found this document to be very difficult to follow. The model was poorly described with many variables not even defined. It was also not clear how the model would be used. It would be impossible for the reader to reproduce the calculations shown in Figures 1-4.

As in every document there was no treatment of goodness of fit or uncertainty in the model.

6. Does the presentation give a description of selected data sources sufficient to allow the reader to form a general view of the quantity, quality and representativeness of data used in the development of emission rates? Are you able to recommend alternate data sources might better allow the model to estimate national or regional default values?

   The data sources for the Tier 2 models are poorly described. They seem to be contained in references 10-12. Were all these data weighted equally for the modeling? How were the data from different studies that had different sulfur contents included in the interpolation? It is not clear which study the paragraph starting with "The study .." refers to. I assume study 12.

7. Is the description of analytic methods and procedures clear and detailed enough to allow the reader to develop an adequate understanding of the steps taken and assumptions made by EPA to develop the model inputs? Are examples selected for tables and figures well-chosen and designed to assist the reader in understanding approaches and methods?

   The model is based on statistical analysis of emission testing performed with gasoline that had two different sulfur levels. The report refers to this analysis as "mixed-model analysis." I am not sure what that means – presumably this is some sort of multivariate model. The chapter needs to describe what the mixed model analysis is. On page 8 the document states that details "can be found in the report." There is not reference provided for this report.

   Presumably the mixed model analysis is used to derive the beta values in equation 17? The report discussing using interpolation for this analysis?

   Equation 17 – This needs to be much better described.

   - What is A? A scaling factor? How is it used? Presumably there are different values of A for different pollutants (e.g. NOx, CO, THC)?
   - I do not understand how the betaS were derived. The text says they were developed by linearly interpolating? However you have many vehicles so presumably you get a whole bunch of betaS values (one for each vehicle tested at the two fuel S levels). In addition the different studies used different fuel S levels? How do you combine the betaS values for different vehicles and different studies? Lumping them together and then

averaging? Presumably the data are stratified by pollutant, model year? What is the uncertainty in these values? How did the values of betaS vary across the vehicle fleet?

- What is listed in Table 2? The BetaS values?

8. Are the methods and procedures employed technically appropriate and reasonable, with respect to the relevant disciplines, including physics, chemistry, engineering, mathematics and statistics? Are you able to suggest or recommend alternate approaches that might better achieve the goal of developing accurate and representative model inputs? In making recommendations please distinguish between cases involving reasonable disagreement in adoption of methods as opposed to cases where you conclude that current methods involve specific technical errors.

   I do not understand the methods or analysis ("mixed model analysis"). This appears to be a purely statistical model as opposed to something based on the underlying physics and chemistry.

   Uncertainty is a key issue that is completely neglected in this chapter. For example, table 1 lists sulfur reduction with 3 significant figures. These values need uncertainty estimates. Uncertainty estimates on these parameters can be derived from the statistical analysis. A better approach would be to challenge the model by performing leave one out cross validation. Ideally both of these approaches would be taken. The complete lack of uncertainty seems like a major weakness of the entire report.

9. In areas where EPA has concluded that applicable data is meager or unavailable, and consequently has made assumptions to frame approaches and arrive at solutions, do you agree that the assumptions made are appropriate and reasonable? If not, and you are so able, please suggest alternative sets of assumptions that might lead to more reasonable or accurate model inputs while allowing a reasonable margin of environmental protection.

   The data seem reasonable. I am not aware of other data.

10. Are the resulting model inputs appropriate, and to the best of your knowledge and experience, reasonably consistent with physical and chemical processes involved in exhaust emissions formation and control? Are the resulting model inputs empirically consistent with the body of data and literature that has come to your attention?

    The chapter presents no data that demonstrates the model provides reasonable results. For example data could be added to Figures 1-4 to help the reader evaluate the model.

Other

Figures 1-4. These appear to summarize the output from the sulfur model. What is the "fuel sulfur adjustment" (which variable, some version of A?)? How is it used? Simply as a scaling parameter on the base emissions? These details need to be clarified.

The review is focusing on the Tier 2 model which applies up to fuel sulfur level of 30 ppmv. It is hard to see the predictions of this model in Figure 1-4 because the x-axis scale goes to 600 ppm. Less than 30 ppmv is less than 5% of this scale. Given the Tier 2 standard for fuel sulfur the report should focus more on the model behavior at current and future sulfur levels (< 30 ppm). For retrospective analyses showing such high fuel sulfur levels may be useful (how long ago were fuel-S levels greater than 400 ppm?). Bottomline is that these figures or a comparable set such focus on performance of the models over the range of current fuel-S levels. Does it even make sense to plot MY 2017 vehicle out at such high fuel S levels?

Although I realize we were not supposed to review the older M6Sulf model, I found the description of the model to be impossible to follow. It is clear that the model is simply a curve fit of the underlying dataset. However, many of the variables in this section are not defined. For example what is A? What do the M6SulfurCoeff values listed on the bottom of page 3 represent? Without more description it is essentially impossible to understand how to apply the model.

It would be useful if this chapter listed the parameterization developed for the M6Sulf model. Presumably these are the wIR, betas's etc. A table defining each variable and listing its value would be very helpful.

You need to define all variables – a short table would be very helpful. What is A2, As,short, As,long, As,Irr ϕ, etc. A is clearly an important symbol. What does it represent? It appears to be some sort of adjustment factor. Is this multiplied with the base emissions to estimate the effects of sulfur? In order for someone to figure out the model, these details need to be much more clearly spelled out.

Equation 8??? I have no idea of the basis for this equation. It is doing some sort of weighting of undefined terms. What is the basis for the irreversibility factor (a sentence to help the reader so that they don't have to look up that grey literature reference).

30 ppmv is the boundary between the two models (Mobile and new Tier 2). Do the two models predict the same effect at 30 ppm? Figures 1-4 suggests that the models link up. What is the basis for the 30 ppm cut – just that it is the tier 2 fuel standard?

Comments on Toxics Report Draft

General comment -- Compared to the other reports there were more typos, broken links, placeholders like "???" in the text, and many typos (e.g. superscripts for references and on numbers, e.g. see Table 47) in this report.

> Does the presentation give a description of selected data sources sufficient to allow the reader to form a general view of the quantity, quality and representativeness of data used in the development of emission rates? Are you able to recommend alternate data sources might better allow the model to estimate national or regional default values?

I thought that the report did not do a good job of providing in text citations to the data sources. Often when the report referred to a data source there was not an in text citation. For example, on page 14 -- they were taken directly from the Complex Model Spreadsheet "*CM Final.xls*". Need a reference for this spreadsheet. This is just on example.

Pre2000 vehicles (Section 2.1) This model is based on old Tier 0 data, which is applied to a large fraction of Tier1 vehicles. There is alot of speciated data for Tier 1 vehicles from the KCVES. Why was a model not developed based on that data? The proposed model should be tested against the KCVES Tier 1 data to demonstrate that it is applicable to those vehicles. At a minimum this needs to be discussed.

> Is the description of analytic methods and procedures clear and detailed enough to allow the reader to develop an adequate understanding of the steps taken and assumptions made by EPA to develop the model inputs? Are examples selected for tables and figures well-chosen and designed to assist the reader in understanding approaches and methods?

The report commonly uses the word "fraction" or "toxic fraction". You need to define fraction of what – VOC, NMOG, THC, etc (presumably each of these is defined using standard EPA definitions). For tables actually defining in header as was done for Table 20 is useful. Also is this a mass or a mole fraction.

Please make sure that all variables are defined – a nomenclature table with units should be added to the report.

Centering data (page 10) – It appears that you are using a different centering approach for older data than for the new model (e.g. eqn 8). Why were different approaches used?

What is meant by model year specific weightings (page 10)? What do these weights represent? Fraction of vehicles for a given year?

Equation 1 – what are the units of the different variables?

Table 8 – Complex model coefficients – these are beta's in equation (1).

Page 13 “For each compound, the model equations as shown in Equation 1, are evaluated for a “base” and a “target” fuel.” This base fuel resides in MOVES? Is this the same as the average fuel listed in Table 7?

Page 14 – equation 3. It was not clear how the weights are being applied. You are trying to derive one adjustment factor for all pre2000 vehicles? Are you driving a separate factor for the 10 different technology classes? This needs to be clarified.

Table 12 -- According to the text these weights represent prevalence for a given technology year. Prevelance means what? Fraction of vehicles based on number, VMT? I am confused that Table 12 lists weights based on “age” as opposed to model years? Is this age relative to 2000? It would be clear to define a base year to calculate age.

Equation 6 -- What is I_{VOC}? Where does the value come from? The standard moves code.

Post2000 organic emissions are based on models derived from the EPAct data. It was not clear if these models are the same as those in the EPAct report. I assumed that they were. If so, the Toxic report needs to specifically acknowledge that. In addition, it should provide specific references to which models are being used as the EPAct report describes a whole bunch of models. Please provide in text citations for the EPAct report.

Table 1 – Are all these hydrocarbons? There are compounds that contain elements other H and C, which I don’t consider to be hydrocarbons.

When you use the term “start” please define it as either cold (e.g. bag 1 of LA92 with appropriate preconditioning) or hot start (bag 3 of LA92).

Page 6 “algorithms” –are these really curve fits as opposed to algorithms?

Page 8 “Toxics inputs for MOVES are not explicitly designed to vary by temperature.” Not sure what this means? The outputs do not vary with temperature? What does temperature refer to? Ambient? Cold versus hot start?

“In addition, while MOBILE6.2 relied on very limited data from heavy-duty gasoline vehicles, MOVES applies Complex Model algorithms to both light-duty and heavy-duty gasoline vehicles” Is there a basis for this extensions. Have additional heavy duty gasoline vehicle data been obtained? If not why is MOVES being extended to heavy duty gas while MOBILE did not?

Page 16 (last sentence of first paragraph) Does MOVES have representative fuel data for different regions and simulations year? Given the focus of fuel dependence of emissions providing the user with a robust set of default fuel values (year and region) would be helpful.

Equation 7 – what is V and what are its units? Equation 7 and associated parameters in Table 13 were derived by fitting MOBILE output. Why not fit directly the original data or use the original parameterization in MOBILE? You claim this equation provides the best fit. What are statistics of fit?

Table 12 -- What do these weight represent? The distribution of different classes of vehicles in different model years? It seems like the minimum age of 2000 vehicle is 13 years (if running a present day simulation).

"It should be noted that the sulfur effects terms in the equations were not included; rather, sulfur effects on toxic emissions were assumed to be proportional to the sulfur impacts on total VOC estimated by MOVES." Sulfur effects in what equations? There is no sulfur in equation 7 (which is the equation that this sentence seems to refer to).

Table 16 exists in Pre-2000 section (2.1– in fact it is in section 2.1.1.1.3) but appears to apply more generally. Move into a more general section of the report?

Do you really want to call ethanol blends gasohol? When I hear gasohol I think of Brazil.

Page 31 "one set representing start emissions and a second set representing hot-running" start emissions is hot start (LA92 bag 3) or cold start (LA92 bag 1, with appropriate conditioning).

There are table reference problems (e.g. see page 32, 35, 38, 40, …). There are other instances of this.

Table 27, 28, etc. Are these parameters from the EPAct report. If so provide citation. Please cite the specific model from the EPAct report, not just the general report.

Page 40 What is OC2.5 VOC?

Page 41—dioxins and furans – "to be similar" You are assuming them to be the same not just similar. Seems like these estimates are very uncertain since they are based on very old vehicles.

The word "data" is plural. E.g. Data were not data was

Diesel PAH data – Similar problems with the partitioning estimates. Partitioning in Schauer study is biased compared to atmosphere. This needs to be explicitly noted in the report. There is a "higher concentration of particles in diesel exhaust" compared to gasoline exhaust in the CVS or plume, but not in the atmosphere. Concentrations in the atmosphere not exhaust is what matters for partitioning.

Table 49 – Particle phase naphthalene? That must be a measurement artifact.

Are the methods and procedures employed technically appropriate and reasonable, with respect to the relevant disciplines, including physics, chemistry, engineering, mathematics and statistics? Are you able to suggest or recommend alternate approaches that might better achieve the goal of developing accurate and representative model inputs? In making recommendations please distinguish between cases involving reasonable disagreement in adoption of methods as opposed to cases where you conclude that current methods involve specific technical errors.

It is not clear why the demarcation for the gasoline vehicles is MY2000 – it seems like the years in which tier 1 or tier 2 vehicles were introduced would make alot more sense. In contrast, the MY2007

distinction for diesel vehicles makes alot more sense than the apparently arbitrary split for gasoline vehicles.

Page 19 section 2.1.1.2 It seems very problematic to be using emissions data from EPAct for a new Tier 2 vehicle to apply to these older vehicles to simulate emissions from high ethanol fuel operations from a pre2000 vehicle. The uncertainty must be very large. Can you run older vehicles on E85? There seems to be little basis for this extrapolation – it seems like you are simply trying to be comprehensive. Ideally a quantitative estimate of uncertainty should be provided for this estimate. At a minimum MOVES should flag the value as massively uncertain.

Phase partitioning of PAH (page 21). This applies to all vehicles (pre2000 and post2000). However it is in the pre2000 section. I found this confusing. Why not have one section that says PAH emissions of all gasoline vehicles estimated using this approach.

More PAH: There is a paragraph that provides the caveat that "gas-particle partitioning of PAHs emission in the atmosphere depends on particle and gas concentrations, exhaust temperature and other factors." It is good to state this. However, presumably the relevant temperature for atmosphere partitioning is atmospheric temperature (not exhaust). This paragraph implies, but does not specifically state, that the gas particle partitioning measured in source test is not representative of atmospheric conditions (or at least not all atmospheric conditions). I think that this caveat needs to be explicitly stated. "The gas particle partitioning of PAHs measured in source tests and implemented in MOVES is likely not representative of atmospheric partitioning."

More PAH: The model use results for composite class, "medium emitters," to estimate gas particle partitioning of all PAHs. Why was a medium-emitters class used? I also suspect that the conditions inside the CVS during the test of these old vehicles (esp. PAH concentrations, PM concentrations, BC concentrations) are not representative of atmospheric conditions (or the newer Tier 2 vehicles). This likely biases phase partitioning towards particle phase. EPA should choose a test in which the conditions concentration and temperature inside the CVS were within the envelope of conditions that likely occur in the atmosphere. This likely would be a test for a cleaner vehicles. An even better approach would be to review the literature of ambient gas-particle partitioning measurements of these compounds and use those values (as opposed to values from a source test). Finally, if the phase partitioning of PAHs is an important output for some of MOVES uses then it is not difficult to implement a gas-particle partitioning model.

Table 20 –The same PAH emissions ratios appear to be applied to all vehicles, which are based on some sort of fleet average from the entire KCVES (or just the pre-2001 vehicles)? It is not clear why this approach was adopted. With this approach you are locking in the emissions based on a fleet that was 10 years old today. How constant were these ratios across the fleet? If they are not constant, why not stratified the emissions into classes (at least Tier1, Tier2) which will allow the model to better forecast future emissions?

In areas where EPA has concluded that applicable data is meager or unavailable, and consequently has made assumptions to frame approaches and arrive at solutions, do you agree that the assumptions made are appropriate and reasonable? If not, and you are so able, please suggest alternative sets of assumptions that might lead to more reasonable or accurate model inputs while allowing a reasonable margin of environmental protection.

In this chapter/report there is wider range of data quality compared to other reports and chapters. Some of the models are based on pretty robust data sources (e.g. basic gaseous organic air toxics), but others are based on data that, at best, are loosely related to the source (Why should fraction of hexavalent chromium emissions from a stationary turbine be representative of onroad vehicles? Or why should emissions from a tier 2 E85 vehicle be representative of emissions from much older vehicle operating on high ethanol blends). I understand the desire for the model to be comprehensive as possible, but the uncertainty of the predictions will vary widely. It does not seem like the model user will have any idea about the quality of the predictions. Ideally each MOVES prediction would provide a quantitative estimate for every prediction. At a minimum the model should provide a grade (e.g. similar to AP42) for each pollutant. For pollutants with robust models, the grade will be high (e.g. A). For less robust models (e.g. hexavalent chromium), the grade would be poor (e.g. F).

In some cases there are important sources of data that have not been utilized (e.g. KCVES to estimate pre2000 vehicle air toxics emissions or PAH emissions for post2000 vehicles).

CNG buses – It seems like there is more data available. WVU has done a bunch of testing on transit buses. Aerodyne research also did a bunch of chase studies of CNG powered transit buses in which they measured high formaldehyde emissions.

Section 2.1.3 Metals –You assume constant emission rates across fleet (which seems plausible, much more so than for PAHs). However, if there were systematic variations in metals emission rates across the fleet why not stratify the model to capture them. What is the quality of the metal emissions? Presumably metal emissions will be sensitive to lube oil therefore it is not clear how widely applicable the data are.

Hexavalent chromium – The speciation is based on stationary combustion turbine testing. Is there any reason to think that is applicable to on-road vehicles? If not, why even report it. At best the results will be highly uncertain. I think this an example of where the model predictions are not supported by robust data.

Page 25 – Why are dioxins and furans expressed as TEQs as opposed to not mass. I am not familiar with dioxins but it struck me as strange. The quality of the dioxins data seemed low.

Are the resulting model inputs appropriate, and to the best of your knowledge and experience, reasonably consistent with physical and chemical processes involved in exhaust emissions formation and control? Are the resulting model inputs empirically consistent with the body of data and literature that has come to your attention?

The report does not provide sufficient information to asses this.

Review of Fuel EffectsTier2Vehicles

1. Does the presentation give a description of selected data sources sufficient to allow the reader to form a general view of the quantity, quality and representativeness of data used in the development of emission rates? Are you able to recommend alternate data sources might better allow the model to estimate national or regional default values?

I think that the presentation of the data sources (specifically test fleet, and fuel composition) could be improved. There is a lot of detailed information in the main EPAct report, which I download and skimmed parts of, but it would helpful for the reader if a bit more information (a few more paragraphs) was provided in the intro about this test program. Here are some examples of the sort of information that would be useful to provide the reader: Were these all relatively new, low-mileage vehicles? What was the variety of emission control technologies? Were the vehicles all port fuel injected? Were all the vehicles 2008 MY? How were the vehicles procured? Recruited from the in-use fleet – if so where? What was the range of each property of the fuels tested in EPAct? What are typical values for each of these properties in actual in-use fuels (summer and winter)?

If all the vehicles were port fuel injected then what is the guidance for gasoline direct injection vehicles which are becoming more prevalent? That seems like the most significant gap in the information.

All of the EPAct vehicles were low mileage, what are the recommendations for higher mileage tier 2 vehicles?

These things seem like important data limitations. Although these issues probably cannot be addressed (these types of vehicles were not in the EPAct test fleet), the document should clearly describe potential limitations of the model so that the reader is aware of them.

2. Is the description of analytic methods and procedures clear and detailed enough to allow the reader to develop an adequate understanding of the steps taken and assumptions made by EPA to develop the model inputs? Are examples selected for tables and figures well-chosen and designed to assist the reader in understanding approaches and methods?

The core statistical model/parameterizations appears to have been derived by the EPAct project and appears to be described in the final report for that project (Assessing the Effect of Five Gasoline Properties on Exhaust Emissions from Light-Duty Vehicles Certified to Tier 2 Standards: Analysis of Data from EPAct Phase 3 (EPAct/V2/E-89) Final Report (EPA-420-R-13-002)). In that (EPAct) report they describe multiple models, but the set of parameters that will be used in MOVES2014 appear to be the same as what is listed in Table ES-1 and ES-2 of the EPAct report (the only exception appears to be the value of the variance listed in Table 3 – why are those different?). This was not clear from reading the fuel effects document. If that is the case (the models were taken directly from the EPAct report), then this document needs to have a short declarative sentencing stating so. "The

models used here were derived and described in the EPAct final report (ref)." Right now the introduction only provides a very qualitative discussion of the EPAct process, but does not explicitly say that the analysis was used here. If the model is different than one of the models derived in the EPAct report then this report needs a lot more discussion of the derivation of the model.

Without reading the EPAct report the reader has essentially no "understanding of the steps taken and assumptions made by EPA to develop the model inputs." The EPAct report is very long and detailed. In addition, they fit multiple models to the data. This chapter would benefit if it provided some more discussion of the EPAct modeling process and why this particular model was chosen (as opposed to one of the other models fit by the EPAct team). This would be a page or so of text. This would give every reader a basic understanding of the model; interested readers could then be referred to the EPAct report for more details. I thought that the air toxics report did a much better job of describing the underlying model(s) than this chapter.

Another shortcoming of this document is that it does not provide some description of the goodness of fit of the model to the original data (part of this should be providing some physical description of what the variance values in Table 3). I skimmed through multiple sections of the EPAct report and could not find that succinctly summarized. A few paragraph (up to a page or two) description of the goodness of fit of the model to EPAct data should be provided as the ability of MOVES2014 to predict fuel effects ultimately depends on the model and how well it describes the data.

Were any exercises performed to test the model with independent data (data not used to fit the model)? Standard techniques such as "leave-one-out" can be used. Alternatively one could use speciated data from other test campaigns to test the model? For example, ARB has extensive data from their surveillance program. This sort of independent evaluation of the model with real world data seems extremely important. This analysis should be performed and described in the report to provide the user confidence in the model.

What was the basis for the assumption "that effects for fuels and temperature are independent and multiplicative."

3. Are the methods and procedures employed technically appropriate and reasonable, with respect to the relevant disciplines, including physics, chemistry, engineering, mathematics and statistics? Are you able to suggest or recommend alternate approaches that might better achieve the goal of developing accurate and representative model inputs? In making recommendations please distinguish between cases involving reasonable disagreement in adoption of methods as opposed to cases where you conclude that current methods involve specific technical errors.

This sort of statistical fitting is commonly done to create "models" to describe fuel effects. The parameters included in the model are known to influence emissions. However, I am not aware of any scientific basis for the underlying mathematical form of the model. If there is one the report would

benefit from a description of it. In addition, without the information on goodness of fit and evaluation of model with independent data as described in the previous section it is impossible to answer these questions.

Beyond a description of the goodness of fit, the major shortcoming of the model is there is no treatment of uncertainty. I would advocate that the model should provide uncertainty estimates (confidence intervals) for every output/prediction. One simple way to provide an estimate would be to use the statistical uncertainty of the fit. This is reasonably straightforward. A more robust approach would also be to try to account for the limitations in the underlying dataset (e.g. lack of GDI). Providing a robust treatment of uncertainty is not easy but it seems essential to ensure that the data are used appropriately. Including uncertainty estimates would be a major upgrade of the model, which may not be possible for this release of MOVES. However, I would strongly encourage EPA to make starting implementing uncertainty a high priority for future releases.

4. In areas where EPA has concluded that applicable data is meager or unavailable, and consequently has made assumptions to frame approaches and arrive at solutions, do you agree that the assumptions made are appropriate and reasonable? If not, and you are so able, please suggest alternative sets of assumptions that might lead to more reasonable or accurate model inputs while allowing a reasonable margin of environmental protection.

   I don't think that there are any statements about data limitations in this section. However, there are some critical data gaps, such as for GDI, higher mileage vehicles, and malfunctioning (gross emitting) Tier2 emitting vehicles. It would be good to at least specifically mention these gaps.

5. Are the resulting model inputs appropriate, and to the best of your knowledge and experience, reasonably consistent with physical and chemical processes involved in exhaust emissions formation and control? Are the resulting model inputs empirically consistent with the body of data and literature that has come to your attention?

This is not covered in this chapter. The trends as report in the EPAct final report seem consistent with expectations.

SO_4 SO_2 Calculator

1. Does the presentation give a description of selected data sources sufficient to allow the reader to form a general view of the quantity, quality and representativeness of data used in the development of emission rates? Are you able to recommend alternate data sources might better allow the model to estimate national or regional default values?

I think that the paper gives a good description of the underlying datasets used to derive the model (in fact I think that these descriptions are better in this document then in some of the other documents).

The models (gas, old diesel, new diesel, CNG) are based on a relatively limited amount of data (one or two studies). The selected studies are relevant because some of them systematically varied key parameters such as fuel sulfur levels (e.g. FUL and DECSE). I am not aware of other studies that have systematically varied these properties.

It seems concerning that some of the core studies (e.g. the KCVES) used gasoline with much higher sulfur content gasoline compare to Tier 2 gas. This means the model has to extrapolate a long way from the reference case. I understand the FUL dataset help do this extrapolation, but it seems strange to have the reference be so far from the current norm on fuel sulfur content.

A major shortcoming of this report is that they show no model evaluation and only limited discussion of goodness of fit. This sort of quality assurance seems essential in an application like MOVES. The model can be evaluated by the many other studies have measured sulfate emissions (e.g. PM characterization by Kleeman group, gasoline component of the gasoline diesel split study, etc.). If some of the parameters are not available in these studies (e.g. fuel sulfur content) the comparison will still provide insight into the suitability of default values. The model should be tested against at least some of these other data to evaluate its robustness. This analysis should be performed and described in the report to provide the user confidence in the model.

2. Is the description of analytic methods and procedures clear and detailed enough to allow the reader to develop an adequate understanding of the steps taken and assumptions made by EPA to develop the model inputs? Are examples selected for tables and figures well-chosen and designed to assist the reader in understanding approaches and methods?

   The basic approach is reasonably well described. I also think that the basic approach of linking sulfate emissions to nonECPM makes sense (and is an improvement from the old approach of linking to fuel S) because it avoids the potentially absurd result if you make assumptions about fuel sulfur content conversion to SO4.

Equation 1 is the core of the model. It was not totally clear how this is implemented in practice. It appears that NonECPM is an output from another part of MOVES2014 and that this model simply scales that fraction using the actual fuel sulfur concentration. Therefore the only independent input into the model is the fuel sulfur concentration (x). All of the rest of the parameters are determined by the reference (listed in Table 1 of main text). If this is the case then it should be clarified in the text.

Presumably there is a default value for this if the user does not know the fuel sulfur content. It would be good to define that value.

There seems to be two assumptions from the It seems like a key assumption is sulfate emission rate from lube oil (SO4o) is fixed for different types of vehicles. Is there evidence to support this assumption? If so it was not adequately discussed in the report. The second assumption is the parameter that describes the conversion of fuel sulfur to sulfate.

I do not understand the treatment of particulate water (Appendix 1 equation 2). Aerosol water depends on the composition of the aerosol and the relative humidity of the exhaust. This can be easily calculated using thermodynamic model such as ISOROPIA. I am not sure how this equation relates to the underlying theory.

A table of variables and definitions would be useful. This is general comment that applies to all chapters.

3. Are the methods and procedures employed technically appropriate and reasonable, with respect to the relevant disciplines, including physics, chemistry, engineering, mathematics and statistics? Are you able to suggest or recommend alternate approaches that might better achieve the goal of developing accurate and representative model inputs? In making recommendations please distinguish between cases involving reasonable disagreement in adoption of methods as opposed to cases where you conclude that current methods involve specific technical errors.

   The model is empirical with the constraint of conservation of mass. This seems like a reasonable approach given the complexity of the system.

4. In areas where EPA has concluded that applicable data is meager or unavailable, and consequently has made assumptions to frame approaches and arrive at solutions, do you agree that the assumptions made are appropriate and reasonable? If not, and you are so able, please suggest alternative sets of assumptions that might lead to more reasonable or accurate model inputs while allowing a reasonable margin of environmental protection.

A limitation that is not discussed is that the sulfur levels of the fuels used in the KCVES are much higher than they are in current Tier 2 gasoline.

Another limitation is the lack of GDI vehicles – as the report states the sulfate emissions depend on sulfur content of the oil/fuel but also combustion conditions. Presumably the differences between combustion in a GDI versus PFI may influence sulfate emission rates.

The major shortcoming of the model is there is no treatment of uncertainty. I would advocate that the model should provide uncertainty estimates (confidence intervals) for every output/prediction. One simple way to provide an estimate would be to use the statistical uncertainty of the fit. This is reasonably straightforward. It appears to have been done in Figures 3-1 and 3-3, which shows the results for the conventional diesel. This needs to be transferred into the core model. Uncertainties should be listed for each of the parameters in Table 1.

A more robust approach would also be to try to account for the limitations in the underlying dataset (e.g. lack of GDI). Providing a robust treatment of uncertainty is not easy but it seems essential to ensure that the data are used appropriately. One way to define this uncertainty would be to challenge the model with additional data that were not used to derive the parameters listed in Table 1. Including uncertainty estimates would be a major upgrade of the model, which may not be possible for this release of MOVES. However, I would strongly encourage EPA to make starting implementing uncertainty a high priority for future releases.

5. Are the resulting model inputs appropriate, and to the best of your knowledge and experience, reasonably consistent with physical and chemical processes involved in exhaust emissions formation and control? Are the resulting model inputs empirically consistent with the body of data and literature that has come to your attention?

There were not sample calculations presented in the chapter. Adding a simple figure that plots sulfate fraction of non-ECPM for a range of reasonable fuel sulfur contents would help the reader understand the model predictions. I suspect that the results will be reasonable a few percent of the PM is sulfate.

Other

Table 1 –

In headers I would add the word “reference” to the last three columns. For example, xB is the reference fuel sulfur level not just the fuel sulfur level.

Table 2-1

Units for sulfur content

Definition of SES variable – sulfur emitted as sulfate suggests that this is ratio or fraction. However this appears to be an absolute emission rate. Why not just call it a sulfate emission rate?

Equations before Table 2-2 – It seems like the Beta1 and Beta2 parameters in this equation are test specific (KC or FUL) and then you make the assumption that they are equivalent.

Table 2-2

Did FUL use FTP or UDDS? In text I thought you said UDDS.

Other comments

Section 2.1 –

It would be good to list the fuel properties that are used in the model (or at least considered in the modeling, since some were dropped out in the analysis) in section 2.1 so that it is clear to the reader what they are. The properties are listed in the intro but it was not clear those were the properties used in the model.

Readers may not know what you mean by second-order and linear terms as these are never defined.

Table 2 –

Units this is % by vol or mass. Same comment for aromatics.

The terms like "etOH x etOH" terms are not defined. Presumably this is the ZZetOHxetOH listed in Equation 3. If so then the table should use the same nomenclature. If not then these terms need to be defined.

The document frequently uses the term "start." Presumably this is actually "cold start" (bag 1) as opposed to "hot start" (bag 3). The term start should always be defined

Section 3. Fuel effect adjustments

It seems like the key here is equation 6 because that is what is actually used by MOVES. You are calculating a scaling factor (equation 6) to apply to the base MOVES emission rate. If that is correct then that should be explicitly stated.

Equations 5 -- I think that it would be useful to list out all the terms.

Equation 6—X (bold) and Beta_in-use are not defined. These are some sort of vector?

Table 3—

It would be helpful if you included a column that had the actual model nomenclature (e.g. ZetOH) as opposed to what you currently list as model terms. Right now the reader may be confused trying to relate the information in Table 3 with the equation (this applies especially to the cross terms).

Why are the variance values in Table 3 different than those in the EPAct report for the same set of model parameters? (This is based on comparing with values in Tables ES-1 and ES2 in EPAct report).

Some discussion of the meaning of the values in Table 3 would be useful to provide the reader some understanding of the actual model. From reading the EPAct report it appears that the sign on the coefficient indicates that it is positively or negatively correlated. The magnitude indicates the size of the dependence?

Section 4—

This table only defines selective values of parameters. It would be useful to have a footnote to a reference where all of the values of each parameter are defined (this would include report and page number).

Table 6

When you write something like ETOHVolume presumably this actual the Z value for this parameter. Should probably try to make this clear in the table in comments column?

Fuel sulfur – ppm volume or mass?

Section 4.1 example

I really like including an example because it can help people understand the model. In this particular chapter, it would be very useful if you actually complete the sample calculation. Provide the reader with a table of input values (actual fuel values and then presumably the Z values for each parameter calculated using the parameters Table 2 -- my understanding is the Z values are what is actually used in the model) and the numerical value of what the model predicts. Having the answer will allow the reader to verify that they understand how to use the model. I would encourage EPA to include this sort of calculation in each of the documents.

There are few places in report where the text is not complete e.g. “add reference to evap report” “Chapter X.X”