Sonoma Technology, Inc.
Innovative Environmental Solutions

September 25, 2017 STI-917053

To: Ira Dassa and Lindsay Kirschner, ICF

CC: Darrell Sonntag and Kent Helmer, EPA

Re: **Comments on EPA's draft report "Population and Activity of On-Road Vehicles in MOVES201X"**

Dear Ira and Lindsay,

I am pleased to be invited to provide review comments on the draft report entitled "Population and Activity of On-Road Vehicles in MOVES201X." Following the instructions in the Peer Review Charge document, I reviewed the entire draft report (Sections 1-17) and Appendix C, Appendix I, Appendix J.3, and Appendix J.4. My review comments are summarized in Attachment A. The required completed COI form is included in Attachment B.

If there are questions, please don't hesitate to contact me at 707.665.9900 or sbai@sonomatech.com.

Sincerely,

Song Bai, PhD, PE
Division Manager, Environmental Modeling
Sonoma Technology, Inc.
1450 N. McDowell Blvd., Suite 200
Petaluma, CA 94954-6503

Attachments

A. Peer review comments
B. Completed COI form

Attachment A. Peer Review Comments

Overall the draft report was very well organized and provided sufficient technical information regarding vehicle population and activity data used in the MOVES default database. My comments are organized into two categories: (a) individual responses to seven specific questions; and (b) additional comments by specific report chapter.

(a) Responses to seven specific questions

1. Do you have any recommendations of better sources or techniques for projecting bus populations and VMT estimates?

The data sources used in this report for estimating bus VMT and population are valid (e.g., the Federal Transit Administration's National Transit Database, the School Bus Fleet Fact Book, and the Highway Statistics), although the data are relatively old. For long-term improvement of collecting bus data, regional transit authority and MPOs (regional planning agencies) would serve as potential sources for getting bottom-up bus population and VMT estimates. Most bus activities are in the large urban areas; therefore, identifying regions and transit agencies with large bus fleets would be a reasonable first step to understand the data availability.

For the approaches described in the report on bus VMT and population projections, I recommend using annual miles per vehicle information to conduct cross-checking. It's good to acknowledge the additional uncertainty for bus VMT and population projections, but the basic assumptions are inconsistent – total heavy-duty VMT growth was used as a surrogate for bus VMT growth (page 22) and total stock growth was used as a surrogate for bus population growth (page 28). Annual miles per bus could be calculated to check whether the heavy-duty VMT growth and total stock growth assumptions are very different. If they are very different (heavy-duty VMT growth vs. total stock growth), it would be important to consider alternative approaches – for example, using total stock growth for bus population projection, and estimating an average (or assuming constant) annual VMT per bus to estimate future year bus VMT growth.

2. Are there any sources of vehicle survivability or scrappage information that are missing, particularly for heavy-duty vehicles? Are there alternatives to this approach for estimating age distributions for future calendar years? As described in Sections 10 and 12, EPA intends to use information from instrumented vehicles to develop default inputs for idle and start activity for heavy-duty trucks. EPA has not yet completed this analysis, so the draft report does not include results. However, EPA would appreciate feedback on its proposed techniques and data sources.

Using registration data as the basis for vehicle survivability and scrappage estimates is typical and appropriate, which is also similar to the approach used by the California Air Resources Board in the EMFAC model for vehicle population and activity modeling. For estimating age distributions, the

approaches look reasonable (e.g., in Section 6) for base age distribution (using 2014 IHS and NTD data), historic age distributions, and projected age distributions. However, the report didn't include discussions about unregistered vehicles (for both light-duty and heavy-duty vehicles). Unregistered vehicles could be short-term or chronic – is there adjustment needed to the IHS or NTD base data to reflect the impact of unregistered vehicles on age distributions (especially for future calendar years)?

Sections 10 and 12 include information on using instrumented vehicles to develop default inputs for idle and start activity. Overall, using instrumented vehicles data (e.g., the Verizon Telematics data and NREL Fleet DNA Database) is a good approach to improve idle and start activity estimates compared to historical approaches. One key issue to address is the regional variation – how idle and start activities vary by state/region and how MOVES adjusts the data processing approach to better reflect the spatial variability in activity data. The limitation of both the Verizon Telematics and NREL Fleet DNA data is that data were collected only for a few states. What is the justification for mapping individual states to the Verizon data for light duty activities (as shown in Figure 10-2)? Should state-specific vehicle population, VMT, and/or other fleet characteristics be considered in mapping individual states to the states with Verizon Telematics data? For heavy-duty trucks, it is very important to develop additional analyses to address representativeness issues in the NREL data (e.g., very high start activities from delivery trucks in Texas and Minnesota, as shown in Figures 12-4 and 12-5).

3. In Section 11, EPA has updated the national default hotelling rate to be consistent with current hours-of-service regulations. For this updated report, EPA evaluated studies of extended activity to inform the hotelling rate, but found that the studies did not report hotelling activity data in sufficient detail for EPA to update its national hotelling rate. For example, Frey et al. (2012) did not report extended idle and APU usage that occurred for stop durations between 3 and 7 hours, and less than 15 minutes. As described in Sections 10 and 12, EPA is currently analyzing truck activity data from the instrumented truck database maintained by the National Renewable Energy Laboratory (NREL). From this data set, EPA can obtain detailed data on extended idling, but not hotelling activity when the main engine is not on, including when the driver is using an APU. By using EPA's current hotelling activity distribution assumptions about the fraction of hotelling that operators idle the main engine, EPA could potentially use the NREL database to inform the national hotelling rate, instead of using the current assumptions with hours-of-service. Would you recommend that EPA use this approach (instrumented truck data on extended idling and assumptions regarding the hotelling activity distribution) to estimate the national hotelling rate?

I would recommend EPA use the instrumented truck data to improve the estimates of the national hotelling rate. Given the calculation approach with equations 16, 17, 18, and xx (see pages 87 and 88), a key assumption is that all trips of long-haul combination trucks are longer than 10 or 11 hours. The NREL data need to be checked (or adjusted as needed) to ensure that this assumption is met; otherwise the hotelling hours could be over-estimated, because any trips shorter than 10 or 11 hours will incorrectly contribute to the hotelling hours calculation. Using VMT for restricted access roads in

both urban and rural areas is a better approach than using just rural VMT, when calculating national hotelling rate. EPA also needs to consider spatial allocation of hotelling activities (e.g., in addition to using state-wide VMT for urban and rural restricted access roads and an overall average hotelling rate).

4. As described in Sections 10 and 12, EPA intends to use information from Verizon Telematics to develop default inputs for idle and start activity for light-duty cars and trucks. Are there any concerns about using this data source? In particular, do you recommend any techniques that would allow us to investigate selection bias or other bias in the data?

As commented for Question #2, using the Verizon Telematics data can potentially improve estimates of default idle and start activity inputs in MOVES. The key issues are related to the data representativeness and temporal and spatial allocations. The QA/QC work described in the report seems reasonable in terms of removing missing or incomplete data. Some specific issues that need to be further discussed include:

a. Justification of individual states mapping (Figure 10-2 default regions for weighting light duty activity).
b. Equation 6 needs to clarify that the regression model handled ordinal categorical variables as independent variables (as shown in Appendix F for the estimated coefficient values).
c. In Figure 10-3, the modeled TIF and Actual TIF are very different for idleRegionID=105 (California), for both passenger cars and passenger trucks; the modeled TIF and Actual TIF appear much more consistent for other regions. This is concerning and needs to be further investigated – is there potential bias in the Verizon Telematics data for California?
d. A typical approach to reduce bias is to apply appropriate weighted average, instead of simple average, when calculating parameters at an aggregate level. For heavy-duty trucks, as described in page 80 (above Section 10.2.2), "The results of each vehicle are then averaged resulting in an even weighting for all vehicles in the sourceTypeID category." – Why is a simple averaging used, not a weighted averaging?

5. EPA has not updated the average speed distributions in MOVES since MOVES2014. New information, such as the telematics analytics used in the CRC A-100 analysis, "Improvement of Default Inputs for MOVES and SMOKE-MOVES," may be available in time for inclusion in the next version of MOVES. How important is it to update the national average speed distributions to account for such data?

In general, speed distributions are important for mobile source emissions modeling. If we are talking about replacing national default data with local data, then speed distributions are unquestionably key inputs and the quality of speed data is very important. In MOVES, emissions are typically more sensitive to the drive schedule and operation mode distribution (especially for county-scale and project-scale modeling runs); this is at a more detailed activity level beyond average speed distributions. For national level application, however, I suggest that EPA consider placing a higher priority on improving the quality of vehicle population and VMT estimates by source type and vehicle

age distribution, instead of average speed distributions. It will be good strategy to first compare speed distributions between the telematics data and the MOVES2014 default, understand how different they are, and then decide whether the average speed distributions need to be updated in MOVES201X with the new telematics data. A potential challenge for using the telematics data would be the representativeness issue – how can we process the telematics data collected from certain states properly so they can represent an overall national average? Much effort will be needed to develop a reasonable approach for repopulating representative speed distributions in the MOVES database (proportions by source type, road type, and hour of day).

6. EPA has not updated the geographical allocation of activity since MOVES2014, but it intends to update these allocations when Version 2 of the 2014 National Emission Inventory is available. Are there any concerns about using the new data with the same MOVES2014 approach described in this report?

Updating the geographical allocation of activity using Version 2 of the 2014 NEI will be an effective improvement for MOVES, and the approach described in the report seems reasonable.

7. EPA is considering a new approach for estimating heavy-duty source mass and heavy-duty fixed mass factors such that they vary by regulatory class and are more closely linked to the actual mass of the heavy-duty vehicles. Details on the proposal for updating source mass values are provided in Attachment A to this Peer Review Charge. Would the new approach be an improvement for MOVES? Are the proposed data sources and analysis approaches appropriate and reasonable? Are there better data sources or techniques for estimating bus source mass?

The new approach will improve the vehicle mass estimates in MOVES. Overall the analysis approaches are appropriate; the vehicle class mapping (Table 1 of Attachment A) is reasonable. For the weighted average gross vehicle mass calculation (see the third paragraph in Attachment A, where multiple FHWA Vehicle Classes are listed for a given source type and regulatory class), VMT is more appropriate than vehicle counts as weighting factors to calculate weighted average mass (because the major emission process is running exhaust that is directly related to VMT, instead of vehicle population). I am not aware of any better data sources for bus source mass. Regional transit authorities or large bus fleet management agencies may be a good start for finding a better data source.

(b) Additional comments by specific report chapter (page number)

The following review comments are more specific regarding the technical contents described in the report. There are minor formatting issues and typos in the report, which are listed at the end of this section.

Section 3.1 (page 20), paragraph above Table 3-1: need to clarify that these revised VMT values ("revised by FHWA in subsequent publications") were used in MOVES201X.

Section 3.2 (page 21), first paragraph: need to clarify the base year in "…applied to the base year HPMS data." Is this base year 2015?

Section 6.2.2 (page 45), first paragraph, last sentence: why was extrapolation used here? Note that the regression approach is typically used to identify statistical relationships within the sample data range and is not appropriate to be used for extrapolation. To be conservative (error to the higher VMT and emissions side), consider the assumption that age 26 and beyond passenger cars have the same VMT as age 25 passenger cars.

Section 6.2.2, Table 6-2 (page 46): need to fill in VMT data for passenger cars; also, why do light trucks of age 29 have high annual miles?

Section 6.2.4 (page 48): mileage accumulation rates #2 "Ages 4 through 16 use…" is very confusing.

Section 9.1 (page 67), last paragraph: this paragraph describes quite a few bus-specific cycles, why does the first sentence say "most of the driving schedules used for buses are borrowed directly from driving schedules used for single-unit trucks"?

Section 9.2 (page 69), first paragraph: "Modeling ramps as part of highway driving using the current driving cycles overstates tailpipe exhaust emissions by less than 3 percent…" What emissions (across all pollutants or for a specific pollutant)?

Section 10 (page 70), first paragraph, last sentence: "… MOVES2014 may not have accounted for the increased amounts of congestion in recent years" – I suggest that EPA clarify idle activities and "stop-and-go" activities in MOVES. Accounting for traffic congestion is related to "stop-and-go" activities and is at least partially addressed in driving cycles. These activities are different from activities in drive-ways, parking lots, or during delivery operations, which are not reflected in any driving cycles. Using Telematics data is potentially addressing both types of activities.

Section 10.2.3, Table 10-7 (page 82): the NREL vehicle population for CA is 47, while the number is 48 in Table 10-5 (see page 79) – need to double check and keep data consistent.

Section 10.2.3, Table 10-8 (page 82): what's the reason behind the CE-CERT data with high total idle fractions for weekend? Some data investigation and quality checking is needed.

Section 10.3 (page 84): the title of this section is for "Work-day…" – why work-day? The description seems to be applicable for all days.

Section 13.3 (page 106): will EPA consider developing HourVMTFraction distributions by source type?

Section 13.4, Table 13-8 (page 110): these starts per day data (e.g., for passenger cars and passenger trucks) are much higher than those presented in Table 12-1 (see page 93). How are the data presented in these two tables related?

Section 13.5 (page 113), first paragraph, last sentence: what is the justification for assuming 60% to account for those unsynchronized trips?

Minor formatting issues and typos:

- Page 23, header of Table 3-4
- Page 25, line 17
- Page 39, last sentence of Section 6.1; second paragraph of Section 6.1.1 "Section 0"
- Page 44, first paragraph of Section 6.2 "Section 4…" and "Section 6.1…"
- Page 77, line 6 of first paragraph
- Page 164, #6, "Replacing the new age 20 and 30+…" should be "…age 29 and 30+"

Attachment B. COI Form

See next page for the completed COI form.

ORGANIZATIONAL CONFLICT OF INTEREST CERTIFICATE

Customer: U.S. Environmental Protection Agency

Contractor: ICF Incorporated, LLC, 9300 Lee Highway, Fairfax, VA 22031

Prime Contract: EP-C-16-020, Work Assignment 0-14

Subject Report: Population and Activity of On-road Vehicles in MOVES201X

Subcontractor/Peer Reviewer: Dr. Song Bai, Sonoma Technology, Inc.

In accordance with EPAAR 1552.209-70 through 1552.209-73, Subcontractor/Consultant certifies to the best of its knowledge and belief that:

✓ No actual or potential conflict of interest exists.

_____ An actual or potential conflict of interest exists. See attached full disclosure.

Subcontractor/Consultant certifies that its personnel who perform work on this contract have been informed of their obligations to report personal and organizational conflicts of interest to Contractor, and Subcontractor/Consultant recognizes its continuing obligation to identify and report any actual or potential organizational conflicts of interest arising during performance under the referenced contract.

Song Bai
Subcontractor/Consultant

9/25/2017
Date