EPA Response to External Peer Review Comments on the EPA Report

Model-extrapolated Estimates of Airborne Lead Concentrations at U.S. Airports

(formerly titled 'Methods for Estimating Airborne Lead Concentrations at Airports Nationwide')

EPA United States Environmental Protection Agency

EPA Response to External Peer Review Comments on the EPA Report

Model-extrapolated Estimates of Airborne Lead Concentrations at U.S. Airports

(formerly titled 'Methods for Estimating Airborne Lead Concentrations at Airports Nationwide')

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

NOTICE

This technical report does not necessarily represent final EPA decisions or positions. It is intended to present technical analysis of issues using data that are currently available. The purpose in the release of such reports is to facilitate the exchange of technical information and to inform the public of technical developments.

EPA United States Environmental Protection Agency

EPA-420-R-20-004
February 2020

Table of Contents

1. Introduction

The United States (US) Environmental Protection Agency (EPA) Office of Transportation and Air Quality (OTAQ) contracted with RTI International to conduct a scientific peer review of a draft technical analysis report that describes the methods for estimating airborne lead concentrations at airports nationwide. RTI International, an independent contractor, facilitated the peer review in compliance with *EPA Science Policy Council Peer Review Handbook, 4th Edition*. RTI selected five peer reviewers with expertise in air quality monitoring and modeling, piston-engine aircraft operations, and potential impacts of piston-engine aircraft sources. Reviewers were charged with evaluating the methodology, assumptions, and supporting data used to estimate concentrations of lead in air from piston-engine aircraft activity at and around airports in the US. Reviewers were also asked to identify any alternative data/approaches that may improve EPA's understanding of the potential impacts of piston-engine aircraft activity on concentrations of lead in air at and near airports. A full description of the peer review process can be found in in Appendix A, which includes the Contractor's report.

1.1 Peer Reviewers

RTI International selected the following individuals to review the report provided by the EPA. Reviewers are referred to by reviewer number throughout the response document, as assigned here alphabetically.

- Reviewer 1: Steven Barrett, Massachusetts Institute of Technology Department of Aeronautics & Astronautics
- Reviewer 2: Michael Kleeman, University of California-Davis Department of Civil and Environmental Engineering
- Reviewer 3: Barbara Morin, Rhode Island Department of Environmental Management
- Reviewer 4: John Pehrson, CDM Smith
- Reviewer 5: Sandy Webb, Environmental Consulting Group LLC

2. Charge to Reviewers

The following charge questions were provided to the reviewers to guide their review and highlight specific areas for input and comment.

1. Sections 1 and 2 describe the nature of how piston-engine aircraft operate for safety and logistical reasons, along with the previous work that EPA and others have conducted to characterize concentrations of lead in air at individual

airports. As stated in the report, conducting detailed air quality modeling or monitoring at all US airports is not feasible due to resource constraints. Please comment on the extent to which this information is clearly described and provide your perspective on the approach selected to utilize modeling from an individual airport in order to characterize concentrations of lead in air at and downwind of maximum impact areas of airports nationwide.

2. Section 3.1 presents the methods to calculate Air Quality Factors (AQFs) at the model airport. Please comment on the approach used to calculate AQFs at the model airport specifically for the purposes of using these factors to estimate concentrations at and downwind of maximum impact areas of airports nationwide.

3. Table 2 and accompanying text in Section 3.2 describe the methods used to estimate piston-engine aircraft landing and take-offs (LTOs) at individual runway ends on a rolling 3-month basis (e.g., apportioning out piston-engine-specific LTOs from total LTOs at each airport, allocating annual activity to daily and then hourly periods). Are these methods clearly described and do you have recommended changes to the steps taken? Please explain any alternative options and provide the location of data sources that would support such alternative options.

4. Section 3.3 presents an analysis to refine estimates of piston-engine aircraft activity using airport-specific data for a subset of airports. Please comment on whether there are alternative airport-specific data, or analysis approaches, that could improve estimates of piston-engine aircraft activity at a subset of airports. In addition, please comment on whether parameters other than piston-engine aircraft activity should be included in analyses to potentially improve model-extrapolated concentrations at a subset of airports, noting that additional parameters are evaluated at all airports in the uncertainty and variability analyses presented in Section 4.

5. EPA provides coarse comparisons of monitored lead concentrations to model-extrapolation results from the national and airport-specific analyses, in Sections 4.1 and 4.2, respectively. In Section 4.3, EPA provides a more detailed comparison of data from lead monitors placed in close proximity to the locations of model-extrapolated concentrations. Please comment on the appropriateness of the approaches to compare model-extrapolated results to monitored concentrations of lead given available monitoring data. Based on your understanding of the methods presented in the report and the comparisons of monitor and model-extrapolated concentrations, please provide your

perspective on the performance of the methods in characterizing the ranges of lead concentrations from piston-engine aircraft at and downwind of US airports.

6. Section 4.3 presents quantitative and qualitative uncertainty analyses of the model-extrapolated results provided in previous sections. Please provide your perspective on the methods used to conduct these uncertainty and variability analyses, as well as the key parameters EPA included in the analyses (based on previous work discussed in Section 2). Please provide your perspective on the application of this analysis to further characterize the range of lead concentrations attributable to piston aircraft activity at airports nationwide.

3. Reponses to Peer Reviewer Comments

The following sections provide the full comments as received from each reviewer along with EPA's response where warranted. Small editorial errors present in reviewer's comments (e.g., misspellings, duplicated words) are corrected in this section; the full, uncorrected comments from reviewers are provided in the contractor's report which is an appendix to this document. References to report section numbers in this document refer to the final report. Full citations to works cited in both this document and the report are available in the References section of the report.

3.1 Response to Comments Received from Reviewer 1: Steven Barrett

Massachusetts Institute of Technology, Department of Aeronautics & Astronautics

Summary Assessment

1. The report represents and in places goes beyond best practice in estimating the range of maximum likely lead concentrations in air due to piston engine aircraft considering over 13,000 US airports.

2. The uncertainty analysis is particularly commendable and provides useful insight on the likely range of concentrations when accounting for biases and uncertainties.

3. The data used is the best available and is appropriately treated in the context of a data-challenged analysis.

4. The model extrapolation method is a rigorously derived approach that is likely to yield very reasonable estimates. While further refinements could be made to account for various factors, these are unlikely to make a material difference and are likely small compared to the uncertainties captured in the uncertainty analysis.

5. Overall this is a comprehensive and high-quality analysis that has been conducted to the highest standard given the limitations of the data available.

6. Notwithstanding this, there are uncertainties in the results which are transparently described and explored both quantitatively and qualitatively.

7. These uncertainties may be significant at any one specific airport of the 13,000, but as a national analysis they are likely to be small. As such the overall conclusions of the analysis are in my judgement robust.

Comments on Sections 1 and 2 (Charge Question 1)

Excerpt from charge question 1: "comment on the extent to which this information is clearly described and provide your perspective on the approach selected to utilize modeling from an individual airport in order to characterize concentrations of lead in air at and downwind of maximum impact areas of airports nationwide."

Reviewer 1 Responses:

8. Section 1 contains a high-quality summary of the nature of aircraft lead emissions, including the quantity and technical purpose of leaded aviation gasoline. Importantly the report also notes that this issue does not apply to jet fuel – which is used in vastly higher quantities.

9. The judgement that the run-up location dominates lead concentration maxima is well justified in my view. This is because aircraft run at high power for an extended period while not moving, cf. takeoff operations where emissions are spread out.

10. In my judgement, EPA are correct in asserting that conducting detailed monitoring and/or modeling at every one of the 13,000 US airports is not reasonable, feasible, or necessary given the aims of the work.

11. The overall approach of a detailed assessment of individual representative model airport is a cogent and well justified method. The notion of an air quality factor to relate emissions to concentrations is rational and transparent, and is a reasonable approach to make assessing airports nationwide tractable problem.

12. EPA make clear that uncertainty and variability characterization is considered in the work, consistent with best scientific practice.

13. Section 1 describes the structure of the report, which is logical and clear.

14. AERMOD is in my judgement a robust and scientifically justified tool for dispersion modeling of the type conducted by EPA. It contains a detailed and practically applicable representation of atmospheric dispersion, and is suitable for application to an airport environment. While it has limitations when applied to jet aircraft sources, these limitations are not an issue when applied to GA sources.

15. The report includes a model evaluation (i.e. relative to data), which is consistent with the very best practice. The R^2 achieved is excellent in the context of atmospheric dispersion modeling, providing confidence in both the methodology used to estimate activity and emissions, and dispersion modeling.

16. The EPA report transparently notes limitations in reproducing modeled values in section 2.2. In my judgement, however, this level of model performance is consistent with the best available approaches and demonstrates a level of model skill that is beyond what I would consider acceptable.

17. My view is that a 7-day monitoring period is more than sufficient to provide confidence in the modeling results and a range of meteorological conditions occurs.

18. The assertion that the 3-month averaging time limits the importance of day-to-day variability is correct and further supports the AQF approach in my assessment.

19. The approaches used for modeling GA aircraft sources are appropriate and consistent with best practice.

20. The approach for estimating aircraft activity and emissions is well described and more than sufficiently detailed.

21. The altitude cut-off and altitude emissions approach is appropriate. While at-altitude emissions do impact surface air quality in my estimation, these are not at all relevant to calculating maximum concentrations as is relevant in this report.

22. The surface and upper air meteorological data stations are close enough to be usefully representative of the meteorology at the airport being modeled. The distance of the surface air station may introduce some uncertainty, but this is likely to be small relative to overall atmospheric dispersion and other modeling uncertainties (except potentially at specific airports).

    EPA Response: We have added this point to the qualitative discussion of meteorological uncertainty in Section 4.4.1.

23. The receptor placements are logical and more than sufficiently resolved.

Comments on Section 3.1 (Charge Question 2)

Excerpt from charge question 2: "comment on the approach used to calculate AQFs at the model airport specifically for the purposes of using these factors to estimate concentrations at and downwind of maximum impact areas of airports nationwide."

Responses:

24. The approach of calculating different AQFs for different types of operation and for single vs. multi-engine aircraft is robust and appropriately accounts for the variability in emissions that is to be expected.

25. The AQFs are logically and clearly defined (e.g. Eq 1). This provides a transparent and practically usable way of characterizing maximum 3-month average lead concentrations given aircraft activity.

26. The use of 14 months is more than sufficient to capture variability in conditions and impacts. It is unlikely this approach leads to any over- or under-estimate that

is of significance relative to the uncertainties inherent in atmospheric dispersion modeling.

27. The specific steps used to calculate AQFs are clearly described and logical.

28. The report correctly notes that the remaining question is the extent to which the results apply to airports throughout the country – addressed in section 4.

29. It would aid clarity if Table 1 used scientific notation (i.e. 1.5×10^{-5} rather than 1.5E-5 etc.).

    <u>EPA Response</u>: Numbers in Table 1 have been changed to scientific notation.

Comments on Section 3.2 and Table 2 (Charge Question 3)

Excerpt from charge question 3: "Table 2 and accompanying text in Section 3.2 describe the methods used to estimate piston- engine aircraft landing and take-offs (LTOs) at individual runway ends… Are these methods clearly described and do you have recommended changes to the steps taken? Please explain any alternative options…"

Responses:

30. The overall approach used to estimate the number of LTOs by piston-engine aircraft is rational and it is hard to see how it could be improved given the available data.

31. The approach is also well-established in NEI use. There will be uncertainties given the data limitations, but I am not aware of alternative reasonably usable data. The additional airport-specific data used serves to understand uncertainties associated with these data limitations.

32. The approach used is clearly described. In particular, Table 2 clearly and in a well-structured way describes the approach along with supporting rationale. (Step 1's title should delete the word "Determine".)

    <u>EPA Response</u>: "Determine" has been removed from the title to Step 1.

33. The 2011 NEI data source is still relevant and appropriate. Although this could potentially be updated and there may be advantages to that, I do not expect this would materially affect results.

> *EPA Response: We agree that the choice of the analysis year is useful to evaluate. Year-on-year changes in the production of avgas (a reflection of national piston-engine aircraft activity) have ranged from a 3% increase to a 13% decrease during the period from 2011 through 2016 (https://www.eia.gov/dnav/pet/pet_pnp_refp2_a_eppv_ypy_mbbl_a.htm). This information supports the conclusion that the choice of a more recent analysis year would not materially affect results of the estimates of ranges of lead concentrations at airports nationwide. We have further addressed this comment in response to Reviewer 1, Comment 50 below.*

34. The use of a national average percentage of GA/AT aircraft that are piston-engined is appropriate given the limited data available, and enables the use of per airport LTO data. This would introduce no uncertainty on average, and some uncertainty per airport. It is unlikely that this uncertainty is significant given the uncertainties inherent in atmospheric dispersion modeling.

35. The division into single and multi-engine aircraft is sufficient to capture the important variability in emissions, along with the division into full LTO and touch and go operations.

36. Using daily activity counts from towered airports to extrapolate activity profiles to other airports is a rational and likely appropriately accurate approach. This includes the use of the closest towered airport for the untowered airports.

37. The wind direction runway assignment approach is appropriate and, on average, is unlikely to introduce significant uncertainty.

38. The assumption in step 12 that the period of maximum activity is assumed to be the period of maximum concentration is robust given that we are considering local passive dispersion modeling. (This assumption could not be transferred to regional chemistry-transport modeling, for example.)

39. The avgas Pb scaling approach is robust and will introduce no uncertainty.

40. Overall these methods are robust and clearly described. There do not appear to me to be viable improvements that should or could be made.

Comments on Sections 3.3 (Charge Question 4)

Excerpt from charge question 4: "comment on whether there are alternative airport-specific data, or analysis approaches, that could improve estimates of piston-engine aircraft activity at a subset of airports. In addition, please comment on whether parameters other than piston-engine aircraft activity should be included in analyses to potentially improve model-extrapolated concentrations at a subset of airports…"

Responses:

41. The criteria used to select airports for further detailed study appears logical and clearly described. In particular, the use of 100% instead of the national fraction of piston-engine AT and GA aircraft is sensible given the potential variability in these numbers across the nation.

42. One possible refinement on a per airport basis that could be applied when evaluating if Pb concentrations come to within 10% of the limit, is to correct for local or nearest available average wind speed where that is known.

    EPA Response: See response to Reviewer 1 comment immediately below.

43. Specifically, the concentration of a passive tracer scales with $\langle u^{-1} \rangle$, where u is wind speed, and angled brackets imply a time average [e.g. Barrett and Britter (2008), Development of algorithms and approximations for rapid operational air quality modelling. Atmospheric Environment 42 (34), pp. 8105-8111. DOI: 10.1016/j.atmosenv.2008.06.020.] If the wind speed at the model airport is v and at a specific airport is u, then the wind-speed corrected concentration would be the concentration estimated by the AQF approach multiplied by $\langle v^{-1} \rangle / \langle u^{-1} \rangle$. This would mean that if the wind speed at a specific airport is lower on average [and so $\langle v^{-1} \rangle$ would be higher] resulting in a higher concentration, this would be captured.

    EPA Response: We have conducted this wind speed correction at all airports as noted in Step 15 of Table 2; results of this wind speed correction and the impact on lead concentrations at the maximum impact site are provided in Section 4.1.

44. This wind speed correction is likely unnecessary in general, but may provide additional robustness in avoiding missing airports that may breach the NAAQS limit. One possibility could be to apply a larger (e.g. 50%) margin instead of 10%, and where the larger margin is reached apply the wind speed correction to determine if the concentrations approach the NAAQS limit.

    EPA Response: See response to Reviewer 1, Comment 47.

45. Such wind speed corrections could be based on nearby or closest ground measurements, or from pre-existing WRF modeling output possibly with appropriate correction for low wind speed conditions.

    EPA Response: We used the same ASOS station wind data used to assign aircraft to specific runway ends. The provenance of the data and the methodology are described in Section 3.2 and Appendix B.

46. I would note that I view this refinement as optional as the approach taken by EPA is scientifically robust. The decision on if to do this is in part a practicality and resource issue. It would, however, provide additional assurance in the result.

    EPA Response: We appreciate the comment and agree that it provides additional assurance to the estimates of lead concentrations provided in this report.

47. Other alternatives to a wind speed correction may also be possible to provide additional assurance that airports breaching the Pb limit are not being missed. For example, further rationale and/or discussion of the 10% as used, or a larger margin.

    EPA Response: We have taken additional steps to identify airports with the potential for lead concentrations at the maximum impact area to be above the level of the NAAQS for lead. See Table 4 for the full description of these steps.

48. The overall approach is in my judgement robust and consistent with best practice, with one optional potential refinement in terms of wind speed correction to account partially for location effects.

Comments on Sections 4.1 and 4.2 (Charge Question 5)

Excerpt from charge question 5: "comment on the appropriateness of the approaches to compare model-extrapolated results to monitored concentrations of lead given available monitoring data. Based on your understanding of the methods presented in the report and the comparisons of monitor and model-extrapolated concentrations, please provide your perspective on the performance of the methods in characterizing the ranges of lead concentrations from piston-engine aircraft at and downwind of US airports."

Responses:

49. While there are inconsistencies (as noted by EPA) both in time and space between monitored data and the modeled data, comparisons are still valuable for assessing confidence in results.

50. The difference in the particular year (2011 vs. other years) is likely to introduce minimal additional error compared to other uncertainties. It may be helpful for EPA to note the change in GA activity (or some equivalent measure, such as avgas sales) nationally over a period of time to give general readers assurance that this is not a significant source of error.

    EPA Response: As noted earlier, EPA agrees that the choice of analysis year is useful to evaluate. Year-on-year changes in the production of avgas (a reflection of national piston-engine aircraft activity) have ranged from a 3% increase to a 13% decrease from 2011 through 2016 (https://www.eia.gov/dnav/pet/pet_pnp_refp2_a_eppv_ypy_mbbl_a.htm).

    Because individual airports may show greater variability than national totals, we evaluated historic operational data from ATADS for the top 50 most active GA airports. We described this information regarding year-on-year variability in Appendix B, and we have referenced the comparison in Section 3.2. The data shows that, while decadal operational trends may be significant, the median year-on-year change in operations at the top 50 GA airports ranges from -4% to 4% for a given year, and the interquartile of airport year-on-year operational changes is between +/- 10% for all years. However, individual airport year-on-year changes range from -25% to +44%.

51. The specific locations being based on judgement will not, in my assessment, introduce any significant error.

52. The approach of providing the monitor location and several model locations graphically, thus showing the complexity of the environments, is a useful way of presenting the data. I believe this to be an appropriate approach given the data available, and is of a very high degree of transparency.

53. My assessment of the coarse monitor/extrapolated model comparisons is that the results are close enough to support the approach taken. It should be noted that the decay rate with distance is significant, and the results are consistent with this.

54. The only outlier in terms of monitored vs. measured results is Airport MM. However, given the mean wind direction it is difficult to compare the monitored data to extrapolated modeled data. Overall the conclusion (5) that the extrapolated modeled data reproduces to a very acceptable degree the monitoring results stands.

> *<u>EPA Response</u>: We have updated the figure captions of the monitor-to-model comparison figures and the text of Section 4.1 and 4.2 to reflect the reviewer's comment that it is difficult to directly compare the monitored data to extrapolated model data for reasons including wind direction, monitor location, and differences in data years for the modeled and monitored data. Further, we acknowledge the reviewer's point that the extrapolated modeled data generally captures expected concentrations evident in monitored results at airports nationwide in accordance with the report's aims; the performance of the model extrapolation in reproducing monitored results may vary from airport to airport depending on local considerations (operations, pilot behavior, active fleet, meteorological conditions, etc.). We expanded our qualitative uncertainty assessment in Section 4.4 to capture this comment. Additionally, we learned after the draft report was provided to peer reviewers that the airport in panel MM conducts the majority of landing and take-off and therefore run-up checks at a different runway and therefore the monitor was not sited to capture the maximum impact site; this figure has been removed from the report.*

55. Section 4.2 contains a thorough and useful discussion of the key factors that result in the variability found. This suggests the reasons are well understood, and that the alternate criteria for selection (100% of aircraft being piston engined) is appropriate.

56. Figure 8 provides a clear depiction of the national and airport-specific results. It may be helpful to also show on this chart the upper bound used in the criteria to select airports for further study.

> *<u>EPA Response</u>: We have taken additional steps to identify airports with the potential for lead concentrations in the maximum impact area to be above the level of the NAAQS for lead (see Table 4). These steps are described in Section 3.3 and the results included in a chart (Table 7) that explains which airports met which criteria. Given these changes and the absence of a single decision metric, we chose not to demarcate a single upper bound value on the figure.*

57. Overall my assessment is that the performance is fit for purpose and, within the limitations of the data currently available, represent a best-practice approach in regulatory modeling.

58. Additional uncertainties are considered in my response to charge question 6.

Comments on Section 4.3 (Charge Question 6)

Excerpt from charge question 6: "Section 4.3 presents quantitative and qualitative uncertainty analyses of the model- extrapolated results provided in previous sections. Please provide your perspective on the methods used to conduct these uncertainty and variability analyses, as well as the key parameters EPA included in the analyses… Provide your perspective on the application of this analysis to further characterize the range of lead concentrations attributable to piston aircraft activity at airports nationwide."

Responses:

59. The inclusion of an uncertainty and variability analysis of this degree of comprehensiveness and quality is to be commended. It is rare to see such an analysis done to this level of quality in regulatory (and academic) practice.

60. Atmospheric stability conditions could also be listed as an uncertainty, along with local roughness conditions, on page 46 end of second paragraph as part of parameter 5.

    EPA Response: This comment has been addressed by mentioning atmospheric stability and surface roughness as additional sources of uncertainty in Section 4.4.

61. In my view the correct key uncertainty parameter groups have been identified and treated appropriately. The choices made have been well justified and explained.

62. The data used to justify the distribution for run-up times appears robust and to materially add value to the work. It is unlikely that assuming the airports for which data is available represent airports more broadly introduces significant error.

63. The same comment (4) applies to avgas lead concentrations.

64. The Monte Carlo approach applied is rigorous and appropriate.

65. Table 5 provides a clear and transparent description of the uncertain parameters and their rationale.

66. In the first row of Table 5, Assumptions column – it is likely more accurate to say that concentrations vary as a power law with distance (as in ~x^{-s}). This does not materially affect the work.

> *<u>EPA Response</u>: We revised the text to note that the lead concentration attributable to run-up decreases as a negative power law with distance from the maximum impact site.*

67. The work supports the finding that the uncertainty in run-up time is key. Given the nature of run-up times (being at the discretion of the pilot in command and being safety critical in nature), it is unlikely more could reasonably be done to estimate and characterize uncertainty in this factor.

68. It may be helpful to give an aggregate expected mean (as a percentage) bias due to the run-up time and avgas lead concentration combined. This could be compared with the mean under-estimate of measured values.

> *<u>EPA Response</u>: We have added aggregate statistics on the median and 97.5th percentiles of the analysis for run-up time and avgas lead concentration to Section 4.3.1 in accordance with the reviewer's suggestion. However, it is difficult to directly compare the results of the Monte Carlo assessment in Section 4.3.1 to the model-to-monitor comparison presented in Section 2.2. The time-in-mode data underlying the model comparison in Section 2.2 was developed from airport-specific recorded time-in-mode survey data; the Monte Carlo assessment time-in-mode distribution is developed from a meta-analysis of time-in-mode data across different airports and different studies. Thus, the uncertainty in the national extrapolation is not necessarily representative of the possible uncertainty or bias in the model airport results.*

69. The comparison in Figure 12 between extrapolated model results (with uncertainty quantified) and monitored data suggests that the baseline modeling and uncertainty analysis captures real-world variability.

70. Section 4.3.2 contains a comprehensive discussion of qualitative factors affecting uncertainty and variability. As well as mixing height, it should also mention atmospheric stability.

> *<u>EPA Response</u>: This comment has been addressed; atmospheric stability is now mentioned along with mixing height in Section 4.4.*

71. It may be helpful to note that rather than the average wind speed, it is the average of one over wind speed, that drives average concentrations.

    <u>EPA Response</u>: We have revised this text to describe the impact of one over wind speed on the average lead concentrations.

72. The sensitivity analysis using varying meteorological factors is useful and provides a clear indication of the potential size of this uncertainty.

73. Overall the application of the uncertainty approach described in the report is robust and provides useful additional information about the uncertainty in model extrapolated values.

74. The results also makes clear that my suggested wind speed correction is indeed optional as this is likely not of great significance relative to the run-up time uncertainty.

3.2 Response to Comments Received from Reviewer 2: Michael Kleeman

University of California-Davis, Department of Civil and Environmental Engineering

Summary:

The purpose of this report is to assess Pb concentrations from aircraft across the United States. A detailed analysis was carried out for a single representative airport using measurements and site specific modeling. Dispersion fields from the representative airport were then extrapolated to the 13,000 US airports using activity data from each specific airport. Based on this generic exercise, a subset of airports were identified where predicted concentrations were close to the Pb NAAQS (or exceeded the NAAQS).

Further site specific modeling was conducted at these target airports to more accurately represent activity data and meteorological conditions.

Comment 1:

The approach summarized above is logical given finite resources, but it would be relatively simple to make improvements at only minor additional cost. For example, it seems possible (likely?) that airports with lower average wind speed than the single model airport used in the current report could have under-predicted Pb concentrations which might mean that some of these airports were not identified for additional analysis. A more-accurate pre-screening could have been performed by pre-sorting the 13,000 airports into approximately 5 categories based on average wind speeds and/or mixing height measured at each location. Detailed modeling could then be conducted for a representative airport within each category using site specific information. The dispersion fields developed for each of these 5 categories could then be extrapolated out to the 13,000 remaining airports by choosing the representative model airport that most closely represents the actual airport. This suggested improvement would more accurately capture the approximate wind speed and mixing height at each target airport. The additional computational expense of this modification would be minor, and it would lead to an improved screening to identify airports that merit more detailed modeling.

> *<u>EPA Response</u>: The approach suggested here for refining the extrapolated lead concentrations is appropriate and logical, yet conducting onsite modeling at additional airports is resource*

intensive (with regard to extrapolating within categories of facilities; we agree with the reviewer that the computational expense of the extrapolation itself would be relatively minor). Based on comments from multiple reviewers, we elected to refine the methodology for extrapolating lead concentrations to account for local wind conditions using a scaler approach. We scaled model-extrapolated concentrations using the relationship between local pollutant concentration and concurrent average inverse wind-speed, as suggested by Reviewer 1 Comment 43, and noted this change in Step 15 of Table 2. The results of this wind speed correction and the impact on concentrations are provided in Section 4.1.

Comment 2:

AERMOD is essentially a steady state plume model that estimates pollutant dispersion based on regional atmospheric conditions measured (or predicted) for the site. AERMOD does not account for complex air flow around buildings or complex air flow generated in the propeller wash region of the aircraft. The report uses AERMOD to predict Pb concentrations at the "maximum impact location immediately adjacent to the run-up area at a runway end". Predicted concentrations at this location are most likely not accurate because the effects of propeller wash on atmospheric mixing have not been accounted for. More complex modeling would be required to accurately predict concentrations at the maximum impact location. This complex modeling should either be performed, or more realistically, the concentrations at the maximum impact location should be removed from the report. This latter option may be preferred since the maximum impact location is generally not accessible to the public and concentrations at this site are therefore not a public health issue. Concentrations should be reported at locations further downwind from the aircraft (25m? or 50m?) where the assumptions inherent in the model are valid.

<u>EPA Response</u>: As noted in Section 2, EPA conducted novel, proof of concept modeling at the Santa Monica airport that has since undergone peer review (Carr et al., 2011). This modeling specifically incorporates initial conditions for aircraft exhaust in parameterizing this source in AERMOD so that maximum impact site concentrations could be evaluated. This work included the incorporation of propeller wash, which creates turbulent mixing over the wings of the aircraft and utilized initial vertical and horizontal mixing using the exhaust temperature and height

relevant for a piston-engine aircraft. Additional information on these parameters is now included in Appendix A of the report. The approach for parameterizing piston-engine aircraft emissions was evaluated using model-to-monitor comparisons at the Santa Monica airport. This modeling framework was then applied to the model airport used in this report and, as described in Section 2.2 of the report, the model performed well at a second airport with regard to estimates of lead concentrations at the maximum impact location.

We have further clarified in the report that the maximum impact site at the model airport was 15 meters behind the aircraft. At airports, including those with high traffic volumes, these locations proximate to piston aircraft exhaust may be in very close proximity (e.g., within 50 meters) to areas accessible by the general public, and are therefore relevant for evaluation in this report.

Carr, E., et al. (2011). "Development and evaluation of an air quality modeling approach to assess near-field impacts of lead emissions from piston-engine aircraft operating on leaded aviation gasoline." Atmospheric Environment **45**(32): 5795-5804.

Comment 3:

Table 1 – remove the column for 0 m based on the issue raised in Comment 2.

EPA Response: See response to Reviewer 2, Comment 2 above. We have changed the text describing what we previously labeled as "0 m" in Table 1 to consistently refer to this location as the maximum impact site. This location was 15 meters behind the run-up location at the model airport as described in Footnote 5 in the report.

Table 4 – remove the column for max site based on the issue raised in Comment 2.

EPA Response: See response to Reviewer 2, Comment 2 on Page 19.

Figure 5 illustrates LTOs associated with Pb concentrations at the max impact site. This should be revised to illustrate LTOs associated with Pb concentrations at some other location where the predicted concentrations are valid. See Comment 2.

EPA Response: See response to Reviewer 2, Comment 2 on Page 19.

Figure 6 illustrates predicted Pb concentrations at the max impact site. This should be revised to illustrate predicted concentrations at some other location where the model is valid. See Comment 2.

EPA Response: See response to Reviewer 2, Comment 2 on Page 19.

Figure 10 should remove predicted concentrations at max impact site as discussed in comment 2.

EPA Response: See response to Reviewer 2, Comment 2 on Page 19.

Figure 11 should plot concentrations at some location other than the max impact site as discussed in comment 2. Suggest choosing location that is well outside zone where propeller wash enhances mixing. See comment 2.

EPA Response: See response to Reviewer 2, Comment 2 on Page 19.

Comment 4:

Table 1 – it is somewhat surprising that the ME concentrations are ~4 times greater than the SE concentrations. Do most ME aircraft have 4 engines? Perhaps this is discussed elsewhere in the report but a note should be made in this table caption to make it obvious to the reader.

EPA Response: ME aircraft typically have two engines, although they can have more than two; the engines used in ME aircraft have greater fuel consumption due to their larger displacement (providing greater horsepower than engines typically used in a single engine aircraft) and typically conduct longer run-up

durations. We have included additional information relevant to this point in Footnote 15.

Comment 5:
Figure 10 caption references shaded blue area but this is not present in the actual figure. Revise caption to match figure.

EPA Response: This comment has been addressed by revising the figure caption (now Figure 12).

Comment 6:

Page 40 discusses the comparison of the predicted vs. measured Pb concentrations at airports where monitoring was performed. The study limits comparisons to locations where the monitor was proximate to the maximum impact area or downwind of that area. This seems to be overly restrictive. The model predicts a continuous field that can be compared to any monitor within a few km of the airport. Model receptor points were arranged in a regular grid and concentrations at sites between those points can be easily interpolated. A comparison should be made to all available measurements at all airports.

EPA Response: The rationale for limiting the comparisons presented to those where monitors were proximate to the maximum impact areas is directly related to the analysis and goal of the estimates being presented in this report. Namely, we are providing estimates of lead concentrations in the maximum impact areas, and therefore present available comparisons of the estimated concentrations with monitored concentrations relevant to this general location. EPA and others have identified the run-up location and downwind areas as the maximum impact locations, due in part to the common attribute of piston-engine aircraft activity conducting run-up in a designated location at each airport. Therefore, maximum concentrations in and downwind of the run-up area is an important commonality among general aviation airports. Aircraft activity outside this common area differs among airports depending on hangar locations, startup and idle locations, taxi-ways, refueling stations and other areas where piston aircraft operate. Because of this, we do not find it instructive for the intended purpose here, to compare modeled estimates of lead

from the model airport to monitored airports where the monitor was located distant from the area of maximum impact.

Comment 7:

The uncertainty in the model predictions should be more fully explored through a comparison between predictions and all available measurements (see Comment 6). All available monitors should be used for this analysis. The uncertainty derived from these calculations more accurately represents the uncertainty of the overall modeling approach than the results of the Monte-Carlo analysis. This uncertainty should be incorporated into the error bars for Pb concentrations at all reported airports.

EPA Response: Please see response to Reviewer 2, Comment 6 above. We disagree that the model-to-monitor comparisons provide a more accurate representation of the uncertainty in the overall modeling approach compared with the Monte-Carlo analysis. First, there has been very limited monitoring at or near maximum impact locations at airports, which prevents the type of analysis suggested. As shown and discussed in the coarse concentration comparisons in Sections 4.1 and 4.2, monitor locations vary both in axial and lateral downwind distance from the maximum impact area. Monitored concentration data and model-extrapolated concentrations also do not necessarily reflect the same operational years. Given that the focus of our analysis is on the maximum impact area and areas downwind (e.g., EPA is not attempting to estimate lead concentrations at all locations on or near airports in the US), and these model-to-monitor comparisons do not account for non-aeronautical sources of lead and variation in background concentration, we are limiting our comparisons of uncertainty to the model parameters that have been demonstrated as being influential at the maximum impact area. Expanding the number of comparisons by evaluating monitors distant from the maximum impact location is not instructive for quantifying concentrations at the maximum impact location or their associated uncertainty.

Further, as noted in the title of Section 4.3, the Monte Carlo analysis is solely addressing the quantitative influence of those parameters that have been demonstrated as being influential in

the maximum impact and downwind areas; it is not meant to be interpreted as capturing all sources of variability and uncertainty. We have text to Section 4.3, to address the reviewer's comment by noting our focus in this section on the key parameters that have been demonstrated in previous studies to impact lead concentrations at and downwind from the maximum impact area at airports while recognizing that additional variables and local considerations may contribute to uncertainty at individual airports. We point the reader to our analysis in Section 4.4 in which we have also expanded the qualitative uncertainty discussion to capture this limitation in response to the reviewer's comment.

Comment 8:

The Monte-Carlo analysis for the effect of run-up time and Pb concentration in fuel seems unnecessarily complicated. Each of these parameters is assumed to be linearly related to ambient concentrations at downwind locations. This has simple and predictable impact on the concentration variable as described below.

Linear relationships between run-up duration and concentrations at various downwind distances are summarized in Table C-1 with the form *concentration=a+b*run-up-time* where *a* and *b* are constants. The linear properties of the expectation operator predict that the variance in the concentration will simply be the variance in the run-up-time multiplied by b^2.

A linear relationship is assumed between fuel Pb concentrations and ambient concentrations with the form *concentration = concentration$_0$ * Fuel-Pb / Fuel-Pb$_0$.* If *concentraton$_0$* is influenced by the variation in run-up time, then we simply substitute this into the equation yielding *concentration = (a+b*run-uptime) * Fuel-Pb / Fuel-Pb$_0$*. The linear properties of expectation should yield the result that the variance in the predicted concentration is simply *b^2 * variance of run-up-time * variance of Fuel-Pb/Fuel-Pb$_0$*.

The simple analysis presented above suggests that the variance of the ambient Pb concentration in response to variance in run-up time and Fuel-Pb can be easily calculated without the need for 10,000 iterations of a Monte-Carlo analysis.

If the actual distribution of Pb concentrations is needed in addition to the calculated variance, then the full Monte-Carlo analysis may be warranted. Figure 11 displays the 2.5% and 97.5% concentrations and so perhaps this motivates the analysis. If these thresholds have some regulatory significance then that information should be described to the reader in the caption for Figure 11 and/or the associated text discussion. If these are arbitrary thresholds designed to show the range of concentrations, then perhaps the standard error (square root of the variance) can be quoted instead at a greatly reduced computational cost.

Suggest consulting with a statistician to confirm the most efficient approach that is still accurate.

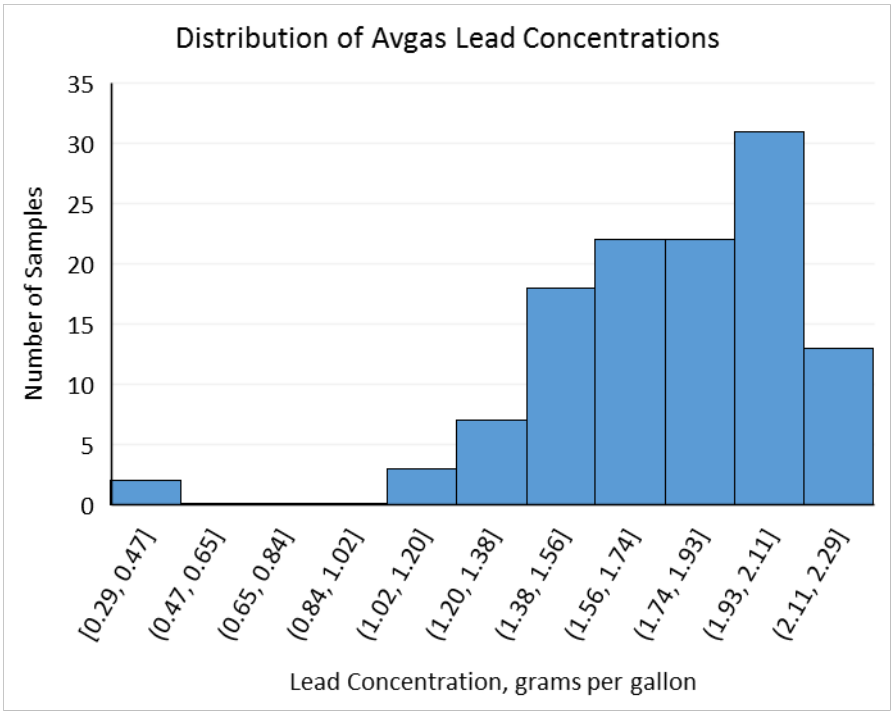
> *<u>EPA Response</u>: Monte Carlo analysis is a useful technique for performing local and global uncertainty analysis, and it can readily be expanded to accommodate additional uncertain parameters. Thus, while the identified parameters could potentially more easily be assessed by techniques like summing uncertainty in quadrature, the focus of this report is in developing a robust methodology for understanding lead concentrations at airports nationwide and associated uncertainty. For that reason, Monte Carlo is an appropriate technique and consistent with uncertainty assessment of other aircraft emissions on impact assessment tools in the literature. Examples from literature include:*
>
> > *Lee, Joosung J., et al. "System for assessing aviation's global emissions (SAGE), part 2: uncertainty assessment." Transportation Research Part D: Transport and Environment 12.6 (2007): 381-395.*
> >
> > *Allaire, D., and K. Willcox. "Surrogate Modeling for Uncertainty Assessment with Application to Aviation Environmental System Models." AIAA journal 48.8 (2010): 1791-1803.*
> >
> > *Simone, Nicholas W., Marc EJ Stettler, and Steven RH Barrett. "Rapid estimation of global civil aviation emissions with uncertainty quantification." Transportation Research Part D: Transport and Environment 25 (2013): 33-41.*

Comment 9:

Page 49 states that “Available data had an average lead concentration of 1.79 g/gal and were normally distributed within the range specified for 100LL (i.e., 1.70 to 2.12 g/gal) (see Appendix C for details on avgas lead data and their distribution)”.
Close inspection of Figure C-1 shows that the measured Pb-fuel concentration is *not* normally distributed but rather is bi-modal with a first peak at 1.55 g Pb / gallon and a second peak at 2.05 g Pb / gallon. This may be an artifact of poorly chosen number of histogram bins. Recommended number of bins would be approximately the square root of the sample size N (~10 in this case). The histogram should be replotted to confirm that it is bimodal.

If the report retains the full Monte-Carlo analysis, the correct distribution for fuel-Pb should be used. If the report drops the Monte-Carlo analysis in favor of just calculating the variance of the ambient Pb concentration as described in Comment 10, then no further action is required.

> *<u>EPA Response</u>: We acknowledge the reviewer’s point that the binning of the histogram was not optimally chosen, and we have replotted the histogram of lead concentrations in fuel in accordance with the reviewer’s suggestion. The replotted histogram, shown below, indeed does not follow a bimodal distribution, and the text of Appendix C has been updated to reflect the improved analysis of fuel lead concentrations.*
>
> *The question of what form and range to select for a “correct distribution” for 3-month average lead concentration at a given facility is a difficult one. Because an airport may be serviced by multiple fuel deliveries over a three-month period; because aircraft landing and taking off at a given airport may have been fueled at a different facility; and because the avgas lead concentration sample data contained noticeable outliers and lacked temporal and spatial resolution, we determined that using the full distribution of avgas lead concentration samples was not appropriate for quantifying uncertainty in three-month average lead concentrations at airports. We use both the central limit theorem and the ASTM fuel specifications to guide the choice of avgas lead concentration distributions used in the Monte Carlo analysis.*



However, we acknowledge the reviewers point that this choice of distribution may be influential with respect to the results of the quantitative uncertainty analysis. Alternative distributions and the inclusion of airport-specific data for both avgas lead concentration and run-up time may be used to better characterize uncertainty attributable to these parameters. While these additional assessments fall outside the scope of this report, we have added text to Section 4.3.1 to address this point.

Comment 10:
The Summary section of the report should reach a more definite final conclusion. Are Pb concentrations at airports a concern or not? Do the report authors believe that the model predictions for NAAQS violations are realistic enough to take action? If this can't be determined based on the current report, what additional actions are needed in order to get to a point where this determination can be made?

<u>EPA Response</u>: We added text to the Summary and Introduction section to clarify the purpose of this report. This report provides ranges of concentrations of lead in air attributable to lead emissions from piston-engine aircraft at US airports and is not

intended to constitute a determination by EPA. As noted in the Federal Lead Action Plan to Reduce Childhood Lead Exposures and Associated Health Impacts (https://www.epa.gov/lead/federal-action-plan-reduce-childhood-lead-exposure), EPA is evaluating aircraft lead emissions and their impact on air quality because this source is currently the dominant contributor to air-related lead emissions in the US. EPA's activities regarding aircraft lead emissions can be found at the following website: https://www.epa.gov/regulations-emissions-vehicles-and-engines/regulations-lead-emissions-aircraft, and the Federal Aviation Administration activities regarding the evaluation of unleaded fuel options can be found here: https://www.faa.gov/about/initiatives/avgas/.

The model-extrapolated estimates of lead concentrations provided in this report cannot be used in making determinations regarding violations of the NAAQS for lead; EPA relies on the lead surveillance monitoring network for such determinations with regard to lead. EPA's guidance on this matter is provided in the National Ambient Air Quality Standards for Lead Final Rule (http://www.gpo.gov/fdsys/pkg/FR-2008-11-12/pdf/E8-25654.pdf).

Comment 11:

Suggest adding the following paragraph (or similar) to the Summary section of the report. A similar statement defining the reasonable scope of the report and proper interpretation of the results should also be included in the introduction.

> "The model predictions described in the current report should be viewed as a screening tool to assess the need for additional measurements of ambient Pb concentrations at airports. The calculated results provide a ranking of the locations where measurements may be most useful to determine if ambient Pb concentrations violate NAAQS levels. The actual maximum concentration values described in this report represent reasonable estimates for ambient Pb concentrations, but they should be verified with measurements before any determination of a NAAQS violation is made."

EPA Response: We agree that additional clarity on this point is needed; we revised the Summary and the Introduction clarifying the purpose of the report and noting that our model-extrapolated values cannot be used to determine compliance with the lead NAAQS.

Minor Comments

Page 34: Sentence reading "Data from previous EPA studies at six airports showed agreement (within 10%) between the number of SE and ME aircraft based at an airport and onsite observations of piston engine aircraft activity the airport (see Appendix B for study details)." should have "at" inserted to read

> "Data from previous EPA studies at six airports showed agreement (within 10%) between the number of SE and ME aircraft based at an airport and onsite observations of piston-engine aircraft activity at the airport (see Appendix B for study details)."

EPA Response: This comment has been addressed.

Page 47: Sentence reading "Run-up emissions accounted for 82% of the 3-month average lead concentration attributable to piston-engine aircraft in EPA air quality modeling at a model facility, and was a primary contributor to emissions in modeling conducted by Feinberg et al. (Section 2, Appendix A){Feinberg, 2016 #11}." Appears to have a reference that was not properly formatted.

EPA Response: This comment has been addressed.

Page 56: Sentence reading "As noted in Section 3.1, the mean of the twelve 3-month average AQFs from the model airport was used to calculate model-extrapolated concentrations; this average was used in order capture the influence on lead concentrations from the full range of wind speeds, mixing heights, and other meteorological parameters that occurred at the model airport." should have "to" inserted to read

> "As noted in Section 3.1, the mean of the twelve 3-month average AQFs from the model airport was used to calculate model-extrapolated concentrations; this average was used in order to capture the influence on lead concentrations from the full range of wind speeds, mixing heights, and other meteorological parameters that occurred at the model airport."

EPA Response: This comment has been addressed.

3.3 Response to Comments Received from Reviewer 3: Barbara Morin

Rhode Island Department of Environmental Management

General Comments-

In general, I am comfortable with the methods used by [EPA] to evaluate impacts at the model airport. I am, however, concerned about the uncertainties associated with extrapolating those results to other airports, as reflected in my comments below. I believe that this analysis appropriately addresses EPA's objective of providing "an understanding of the potential range in lead concentrations in air at the approximately 13,000 airports with piston-engine aircraft activity in the US." However, if modeling results are to be used, alone or in conjunction with monitoring results, to demonstrate compliance with the NAAQS or to evaluate the potential for site-specific exposures, it is my hope that those analyses will utilize airport-specific information rather than relying on the AQFs derived in this study.

> *EPA Response: We appreciate the Reviewer's feedback that the analysis included in the report appropriately addresses the objective described in the report. We incorporated additional text in the Introduction to the Report to further emphasize that in making determinations regarding violations of the NAAQS for lead, EPA relies solely on the lead surveillance monitoring network. EPA's guidance on this matter is provided in the National Ambient Air Quality Standards for Lead Final Rule (http://www.gpo.gov/fdsys/pkg/FR-2008-11-12/pdf/E8-25654.pdf).*

Charge Questions and Responses-

1. Sections 1 and 2 describe the nature of how piston-engine aircraft operate for safety and logistical reasons, along with the previous work that EPA and others have conducted to characterize concentrations of lead in air at individual airports. As stated in the report, conducting detailed air quality modeling or monitoring at all US airports is not feasible due to resource constraints. Please comment on the extent to which this information is clearly described and provide your perspective on the approach selected to utilize modeling from an individual airport in order to characterize concentrations of lead in air at and downwind of maximum impact areas of airports nationwide.

   While the information presented in Section 1 and 2 is clear and useful, short (one or two sentence) explanations addressing the following topics would further aid in the understandability of this material:

- The document explains the function of lead in aviation gasoline. Since lead previously served a similar function in automobile gasoline and has been banned from that fuel for more than 30 years, a sentence about why removing lead from aviation gasoline has not yet been considered feasible would be helpful.

  EPA Response: *We have added a footnote that unleaded motor vehicle fuel cannot generally be used in piston-engine aircraft because of the minimum octane requirements as well as other carefully controlled fuel parameters in avgas.*
  It is relevant to note here that the general aviation industry and fuel providers, together with the Federal Aviation Administration are currently engaged in a multi-year program to identify and evaluate unleaded fuels for use in piston-engine aircraft (information available at the following link: http://www.faa.gov/about/initiatives/avgas/*).*

- It took me a little while to reconcile the statement that "Run-up activity is estimated to contribute over 80% of the lead concentrations at and downwind of the area where the runup mode of operation occurs" with the emissions breakdown by operation type in Table A-1, which associates 36% of ME LTO emissions and 15% of SE LTO emissions with run-up operations and Figure A-9, which shows fuel consumption rates during run-up to be similar to those during approach and less than those during take-off and climb modes. I assume that this is because the aircraft is at ground level and stationary during run-up, unlike during taxiing and take-off operations, as well as the fact that run-up operations take place near the end of the runway, but further explanation would be helpful.

  EPA Response: We have added text to this section to clarify the contribution of run-up emissions to lead concentrations at and immediately downwind of the run-up location. The text we added communicates that run-up operations are typically conducted adjacent to the runway end from which aircraft take-off and the brakes are engaged so the aircraft is stationary. As a result of the stationary aircraft, duration of run-up, and high fuel consumption rate, emissions from run-up activity are the largest contributor to local maximum atmospheric lead concentrations; run-up emissions are estimated to contribute over 80% of the lead concentrations at and immediately downwind of the area where the run-up mode of operation occurs, even though this mode of operation does not have the highest fuel consumption rate. We refer the reader to Appendix A of the report for more intormation.

- A sentence about why RHV was chosen as the model airport would be useful.

  <u>EPA Response:</u> EPA selected RHV as the model airport because it was representative of general aviation airports where piston-engine aircraft operate. Additionally, this airport allowed EPA to evaluate the use of AERMOD in a more complex airport setting (e.g., parallel runways) compared with the earlier proof of concept modeling EPA conducted at SMO. We have added a sentence to the report noting this information.

- Is longer-term monitoring being conducted at RHV to evaluate annual impact predictions?

  <u>EPA Response:</u> Lead monitoring at the RHV airport is required to continue per requirements stipulated in the NAAQS for lead (lead concentrations measured above half the NAAQS necessitate continued monitoring). These data could be used to understand year-to-year changes in lead concentrations at this facility.

- Although I understand that the short-term model-monitor comparisons were considered to be within the acceptable bounds, the fact that the model under-predicted the monitored values on 6 of the 7 days at the maximum impact site and on all 7 days at the downwind site is not reassuring. Was any consideration given to adjusting model results to account for this under-prediction? An explanation of this decision would be helpful.

  <u>EPA Response</u>: We acknowledge that our model airport evaluation suggests the modeled concentrations are somewhat lower than the monitored values at the model airport. As stated in the report and noted in this comment, the difference we observed between modeled and monitored concentrations is within the commonly held acceptable bounds; moreover, when the modeling approach was applied elsewhere, results showed under- and overestimates (Carr et al., 2011; Feinberg et al., 2016). One could conduct a sensitivity analysis to evaluate the impact of adjusting model results for under-prediction (e.g., apply corrections based on the seven days of model-to-monitor comparison at the model airport). We did not conduct such a sensitivity analysis, in part because it would presume that the difference between modeled and monitored concentrations quantitatively captures all of the relevant uncertainties and is consistently, directionally correct. In addition, this type of sensitivity analysis does not account for any uncertainty in monitoring results, which may contribute, along with model uncertainty or bias, to

differences modeled versus monitored concentrations. Rather than adjust modeling results at the model airport a priori, we utilized a series of uncertainty analyses to evaluate how key parameters may impact model-extrapolated concentrations at airports nationwide. The specific uncertainty analyses were selected based on observations at several airports which identified key parameters that impact modeled concentrations of lead from piston-engine aircraft activity. Each uncertainty or sensitivity analysis is described in detail in Sections 3 and 4, with supporting information available in Appendix C.

Carr, E., et al. (2011). "Development and evaluation of an air quality modeling approach to assess near-field impacts of lead emissions from piston-engine aircraft operating on leaded aviation gasoline." Atmospheric Environment **45**(32): 5795-5804.

- The concept of aircraft emissions as volume sources (page 10) is counterintuitive, since emissions are from the tailpipe. This topic is explained further in A.1.5, but it would be helpful to either add a sentence to the introduction to address that issue or to add a reference to A.1.5 in the introduction.

  EPA Response: We added a reference to Appendix A, Section A.1.5.

- Page 10 states that "wind direction data were used to identify the runway end from which piston-engine aircraft took off during each hour of each day in the year of modeling." To facilitate analysis of monitoring data around TF Green Airport, RIAC, the operator of that airport, provides RIDEM/RIDOH with data on the time, aircraft and runway for each LTO. Are similar data available for RHV or for any of the other airports evaluated? If so, was there any attempt to use such data, even if incomplete, to verify the runway assumptions based on wind direction?

  EPA Response: EPA or EPA contractors visited ten general aviation airports (RHV, SMO, DAB, ACK, CRQ, SQL, DAB, MRI, PTK, and VNY) to visually verify the use of wind direction to evaluate the active runway used. We do not have runway-specific identification of activity for 2011 from these facilities (other than RHV) to compare with the runway activity estimated in this analysis. We conducted several sensitivity analyses to evaluate areas of uncertainty such as runway assignment during the peak

period of piston-engine aircraft activity. These additional analyses are described in Section 3.3 with results presented in Section 4.2.

2. Section 3.1 presents the methods to calculate Air Quality Factors (AQFs) at the model airport. Please comment on the approach used to calculate AQFs at the model airport specifically for the purposes of using these factors to estimate concentrations at and downwind of maximum impact areas of airports nationwide.

   - I would appreciate a sentence of phrase explaining the purpose of touch and go operations. I assume that these are training exercises.

     <u>EPA Response</u>: We have added additional text to Footnote 3 on touch-and-go operations to explain that they are part of pilot training.

   - Page 12 characterizes the maximum impact site as "the runway end at which LTOs most frequently occurred at the model airport facility." The document says earlier that 80% of the lead concentration at the maximum impact site is from run-up operations. I understand that run-up operations are generally conducted at the runway end that is being used for take-offs. However, since run-ups are not associated with landing operations, I assume that the number of take-off operations, rather than the number of landing and take-off operations, would be the relevant factor for determining the maximum impact location. This may make a difference at some airports, if the diurnal patterns for landing and take-off operations differ.

     <u>EPA Response</u>: While run-up is the main contributor to lead concentrations at the maximum impact site, both landings and takeoffs contribute to lead concentrations at and downwind of the maximum impact location. Since, over a 3-month period, each takeoff must also be associated with a landing, and airports do not identify operations as take-off vs. landing, but just 'operations', it is appropriate to model each landing and takeoff as a linked pair. While for each particular landing and takeoff pair, the takeoff must precede the landing, because general aviation operations are typically of short duration, we do not expect there to be a significant difference between the landing profile and takeoff profile across the day other than at the margins. Given that it is significantly less computationally intensive to run the extrapolation model for each LTO cycle as a pair and because we have limited evidence of any significant difference between the landing and takeoff profiles across the day, we model the two as a single operational unit (an LTO or a T&G). We have added to Appendix B a short discussion of the landing and takeoff profiles, showing that the landing and takeoff profiles at six airports are

not appreciably and consistently different across days where aircraft takeoffs and landings were surveyed. Further, we discuss that given evidence from modeling, literature, and surveys of aircraft operations, the 3-month average lead concentrations are not expected to be sensitive to diurnal profile uncertainty. We show that using a generic diurnal profile is expected to contribute only 2% uncertainty to lead concentrations, based on a comparison of profiles at airports where both landing and takeoff survey data exist. However, we acknowledge that diurnal profile may contribute additional uncertainty at individual airports if there are specific local operational patterns that would significantly separate take-off operations from landing operations diurnally. A discussion of diurnal profile uncertainty has also been added to Section 4.4.3 with additional details in Appendix B.

- I am also a little confused that there is no distinction between landing and take-off in calculation of the AQF ratio (i.e. that the equation used to calculate AQF uses LTOs), since emissions during landing and take-off operations are very different (and take-offs also involve run-up operations). Is the assumption that every landing is associated with a take-off? Does the fact that those activities may take place on different runways (e.g. because of diurnal wind variations) impact the accuracy of this calculation?

  <u>EPA Response</u>: As noted above, EPA does make the logical assumption that each takeoff is associated with a landing at a given airport. While some airports may have a significant fraction of itinerant flying (i.e., from one airport to another), we make the simplifying assumption that over a three-month period, modeling each take-off as being associated with a landing is appropriate.

  It is not expected that the diurnal profile in take-offs and landings is a significant parameter that would impact rolling three-month average lead concentrations at most airports. Previous modeling of aircraft lead concentrations found that monthly atmospheric lead concentration estimates showed relatively small variations based on a wide range of input diurnal profiles because any hourly differences (from, for example, wind speed) average out over longer periods [Feinberg and Turner, 2013]. Extending the averaging period from one month to three months should further reduce modeled concentration sensitivity to diurnal profile.

  To further understand sensitivity to diurnal profile, EPA examined the impact of using a generic "operational" diurnal profile vs. a "landing" or

"take-off" specific diurnal profile as suggested by the reviewer's comment. For this choice of diurnal profile to be impactful on three-month averaged concentrations, two factors would need to occur: the difference between the diurnal profile of landings and the diurnal profile of take-off would need to be significantly different, and the average wind direction at the time of a potential overestimate of take-offs would need to be significantly different than the average wind direction at the time of a potential under-estimate of take-offs.

EPA examined the landing and take-off patterns at RHV (the model airport) and at five airports for which airport survey data was available. Given that piston aircraft do not typically operate at night and that an aircraft must first take-off for it to land, there is an expectation that take-offs will (on average) occur earlier than landings. However, piston-engine aircraft typically perform short operational missions. Thus, while at the margins landings should occur later than takeoffs, we do not expect the profile of landings and takeoffs to differ significantly.

Airport surveys at each airport reported counts of landing and take-offs during operating hours or a subset of hours for between three and six days of operation. Operational survey data were excluded for any day that did not have survey data covering at least 80% of operational hours or for any days where both landing and take-off data were not available. The figure below shows the difference in percentage points of the landing and take-off diurnal profiles at each of these airports as reported in survey data. The data confirms the expectation that take-offs were relatively more prevalent than landings over the first hour of monitored operations (and, consequently, in the last hour of operations, take-offs were relatively less prevalent than landings), but that profiles were otherwise similar over the day. The figure shows that, on average the difference between a landing diurnal profile and a takeoff diurnal profile is 2.6 percentage points. Further, 95% of examined hours show a difference of less than 6 percentage points between the landing diurnal profile and the take-off diurnal profile. Thus, using a generic "operation (LTO)" profile will, on average, over or under predict takeoffs by 1.3% in any given hour.

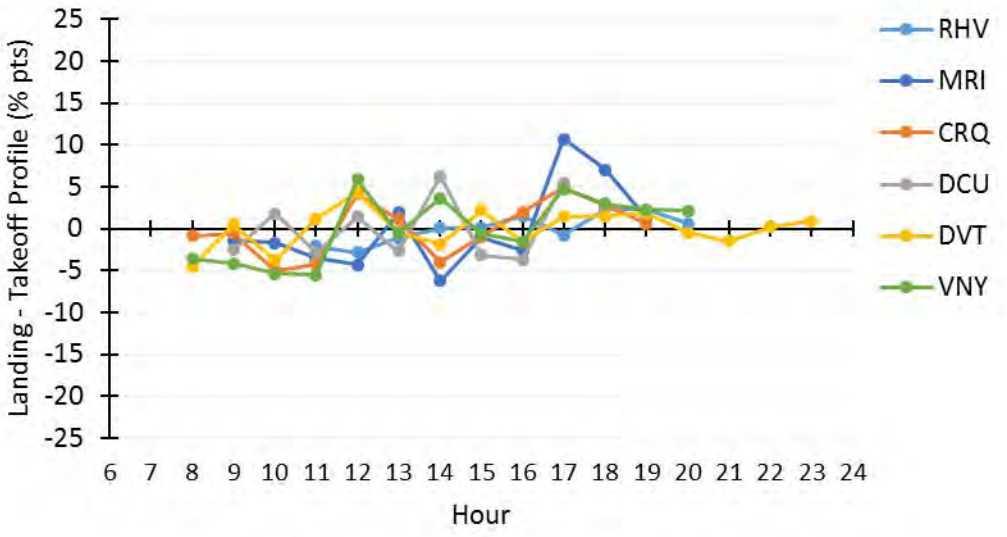


Figure 1: Difference between landing diurnal profile and takeoff diurnal profile from 3 to 6 days of survey data at each of 6 airports.

In our analysis of individual airports, we consider +/- 10% maximum 3-month concentrations (in addition to considering variation in other more sensitive parameters such as the expected split between multi- and single-engine aircraft), which far exceeds the +/- 1.3% uncertainty from differences in landing and take-off diurnal profile.

Further, while the average difference between a generic operation diurnal profile and a take-off only diurnal profile is 1.3%, the actual uncertainty in resulting 3-month average concentration may be even smaller. Since aircraft are assigned to runway primarily based on wind direction, a modeled difference in operations would require the wind direction to change significantly from the time in which takeoffs may potentially be overestimated to the time in which they may potentially be underestimated.

The figure below shows the average wind direction at 938 ASOS stations nationwide for each hour of the day. Wind direction is normalized such that the wind direction at 00:00 is 0° at all stations. Each ASOS station is plotted along a circle of different unit radius, and each hour is modeled by a dot where angle represents difference in wind direction from 00:00 and color represents time of day. The figure shows that across all ASOS stations, 86% of all hours have average wind direction that fall within 90° of the initial recorded wind direction. Therefore, for example, at a single runway airport, even if the diurnal profile were moving 2% of operations

from the morning to the afternoon, there is an expectation that, averaged over a three-month period, those operations would still generally be assigned to the same runway end.

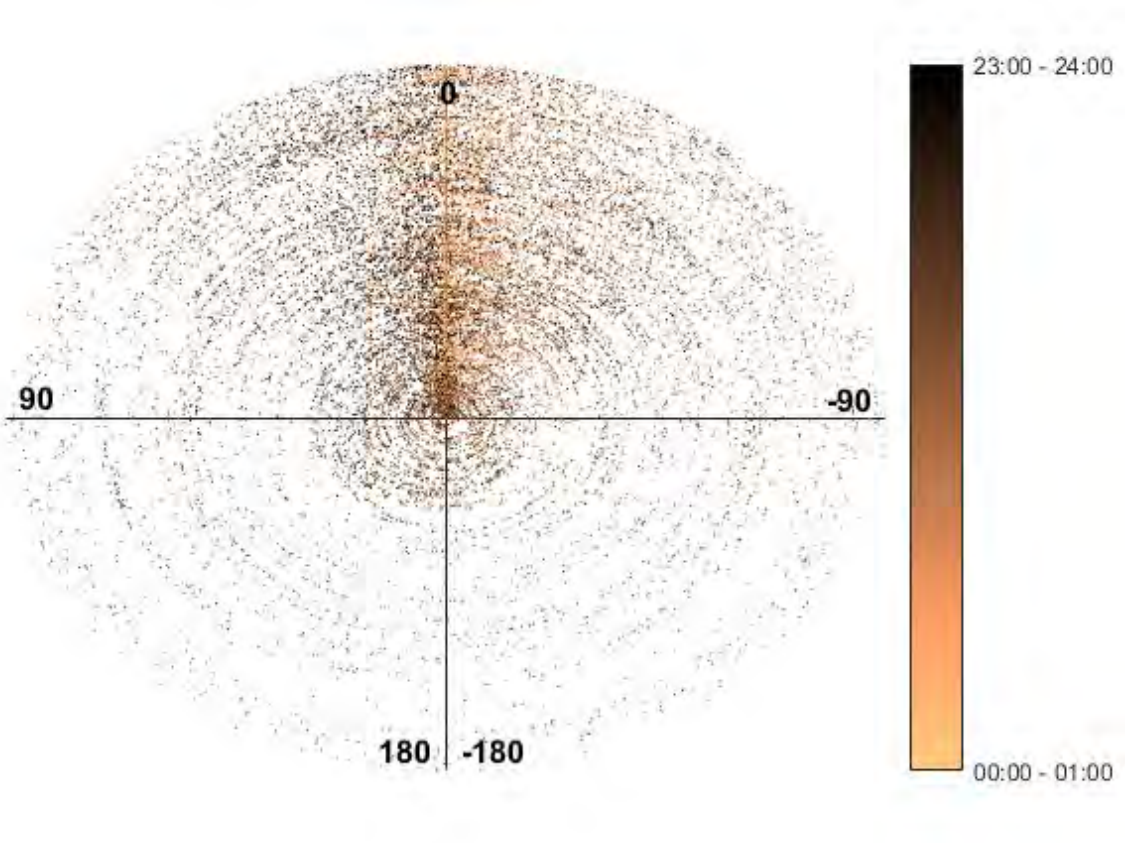


Figure 2: Normalized average wind direction by hour of the day at 938 ASOS stations.

Given the evidence that 3-month average concentrations are not highly sensitive to diurnal profile, the selection of the RHV diurnal profile is appropriate for the national analysis of lead concentrations. To better explain and document the modeling of diurnal profile, we have added the above discussion to Appendix B to support the existing discussion of diurnal profile. We have referenced this discussion and the associated uncertainty from using a single diurnal profile in an expanded section on Operational Parameter uncertainty (Section 4.4.3).

Feinberg, Stephen, and Jay Turner. "Dispersion Modeling of Lead Emissions from Piston Engine Aircraft at General Aviation

Facilities." *Transportation Research Record: Journal of the Transportation Research Board* 2325 (2013): 34-42.

- The document states that "AQFs are calculated as the ratio of the average lead concentration over rolling 3-month time periods to piston-engine aircraft LTOs at the most frequently used runway end over the same 3-month period." Are there cases where a receptor is impacted by LTOs from more than one runway? This is the case at TF Green. While this may not have a significant impact at the maximum receptor, it can be a factor in calculating three-month average concentrations at downwind receptors, so focusing only on one runway may underestimate those impacts.

  <u>EPA Response</u>: EPA agrees that the approach described in this report would underestimate impacts for cases where the maximum impact site is impacted by piston-engine aircraft activity in addition to that characterized at the model airport (i.e., multiple runways, additional taxi and idle locations). An airport-specific assessment would be needed to comprehensively evaluate lead emissions and concentrations for complex airports in which a range of aircraft activities could be influencing lead concentrations downwind from the maximum impact area.

- The document states that "In order to average across the largest range in meteorology inputs to AQFs (e.g., wind speed), the resulting 12 AQFs were averaged to provide a single 3-month AQF for each aircraft class, operation-type, and location combination." Since the NAAQS is a not-to exceed, maximum 3-month concentration, why would you average the 3-month AQFs calculated? Due to seasonal variations in meteorology and emissions, the 3-month AQFs calculated for some 3-month time periods would legitimately be higher than for other times of the year. By averaging the 12 values, you are losing that range, and potentially underestimating 3-month averages that may exceed the NAAQS.

  <u>EPA Response</u>: We acknowledge that the use of the maximum AQF from the model airport would provide higher estimates of extrapolated lead concentrations; as noted in the report, the maximum weighted AQF is 23% higher than the average weighted AQF. To achieve our goal of using the model airport to best characterize lead concentrations at airports nationwide and in acknowledgement of the importance of meteorology on lead concentrations, we used the average AQF and then wind-speed corrected the estimated concentrations as a means to

further refine and make comparisons with the lead NAAQS (the wind-speed correction is a new analysis added in response to peer review comments we received).

3. Table 2 and accompanying text in Section 3.2 describe the methods used to estimate piston- engine aircraft landing and take-offs (LTOs) at individual runway ends on a rolling 3-month basis (e.g., apportioning out piston-engine-specific LTOs from total LTOs at each airport, allocating annual activity to daily and then hourly periods). Are these methods clearly described and do you have recommended changes to the steps taken? Please explain any alternative options and provide the location of data sources that would support such alternative options.

   The methods are clearly described. However, while I understand the need to generalize when extrapolating to a large number of other airports, I question whether there is so much uncertainty associated with those extrapolations that the predictions are not meaningful. For instance:

   <u>EPA Response</u>: We address each of the Reviewer's specific points below, but address the general concern regarding uncertainty in the extrapolated concentrations here. As noted in the introduction statement, the methods used are intended to provide results that are informative for understanding ranges of lead concentrations in air at airports nationwide. As such, we provide a detailed analysis of uncertainty and variability in key parameters in Section 4 of the report and where individual airports are evaluated, we utilize airport-specific data to the fullest extent possible, while maintaining a quantitative uncertainty analysis of lead concentrations at these facilities as well. We recognize the reviewers concern that use of these model-extrapolated concentrations outside the aims if this report may be less meaningful. We have explicitly added a statement that these results are not to be used to determine NAAQS attainment status in the Introduction and have revised the report to be more specific as to the interpretation of the results.

   - Appendix B says "A comparison of the diurnal profiles across these four facilities (airports) shows the same basic features: a ramp-up of activity in early morning, peaks in activity in late morning and early afternoon, and decreasing operations in the evening." However, the diurnal patterns for RHV shown in Figure B-2 appear to be distinctively different from those for the other airports. At RHV, the volume of operations remains high through most of the day, peaking at 4:00 PM on weekdays, while, at the other airports, activity peaks in the morning and drops off in the afternoon. This is significant because atmospheric dispersion characteristics tend to be much less favorable in the

early morning than at midday or in the afternoon, so emissions from morning flights could have significantly higher impacts than emissions that occur later in the day. As an example of this phenomenon, measured pollutant levels near highways tend to be significantly higher during morning rush hour than during evening rush hour, despite similar traffic volumes and congestion levels. We see similar effects at monitors near TF Green Airport.

EPA Response: We agree with the reviewer that increased atmospheric stability (lower mixing height) in the morning hours generally leads to greater concentrations from ground-based emissions compared with afternoon hours. Wind speed is also strongly related to mixing height, as shown in Figure 14, with a correlation coefficient of 0.98 at the model airport; we have conducted a new analysis of the impact of wind speed on lead concentrations which refines the extrapolated estimates of lead concentration (See Section 3.2 and Step 15 of Table 2). With regard to the impact of the diurnal profile in aircraft activity on the extrapolated concentrations, it is useful to point out that while one can observe potential differences between airport diurnal profiles presented in Figure B-2, the diurnal profiles across each of the airports are broadly consistent. For example, while surveys at RHV show 4.5% more operations at 16:00 relative to RVS (normalizing for total operations), the surveys also show RVS having 4.5% more operations at 18:00 relative to RHV. Further, as discussed in the response to comments above, while hourly concentrations may be sensitive to operational profiles, long-term average concentrations (such as the 3-month average concentrations developed here) are not expected to be sensitive to diurnal profile. We added discussion to Appendix B and Section 4.4.3 to further discuss uncertainty contributions of diurnal profile modeling choices. Lastly, it is relevant to note that differences in the diurnal profile across the four airports for which extended operational survey data are available may be driven by local or regional operational patterns, seasonal differences, or an artifact of survey length and methodology, which, when taken together, further suggest that the assumption that these activity profiles are broadly similar is a reasonable one given available evidence.

- As mentioned above, impact calculations were done only for the most used runway end. There are several reasons why this may underestimate impacts. First, some downwind sites may be impacted by LTOs from more than one runway end, so discounting all flights except those that on the most-used runway may underestimate 3-month average exposures at those locations. In addition, the number of flights isn't the only determinant of impact. If wind

speeds associated with one wind direction are stronger than those associated with another direction (not a far-fetched idea), impacts associated with a runway end with lighter average wind speeds may be higher than impacts from a somewhat more well-used runway end that is associated with higher wind speeds. For evaluating impacts on ambient air, as would be done to determine compliance with the NAAQS, the proximity of the runway end to the property line would also be an important factor.

> *EPA Response: We acknowledge that the methods used in this assessment do not include potential impacts at specific airport facilities where activity at multiple runways or taxiways may increase lead concentrations at downwind receptors. We also acknowledge the important impact of wind speed on ambient lead concentrations, and have incorporated a new analysis to refine the extrapolated lead concentrations by correcting for wind speed at each individual airport during the period of maximum activity. For the analysis of specific airports where estimated lead concentrations were above the level of the lead NAAQS, we used visual inspection of Google Earth images and on-site inspections to evaluate whether the area within 50 meters of the maximum impact site has unrestricted access. As noted in responses to comments above, and in the report, EPA relies solely on the lead surveillance monitoring network for lead NAAQS attainment determinations. EPA's guidance on this matter is provided in the National Ambient Air Quality Standards for Lead Final Rule (http://www.gpo.gov/fdsys/pkg/FR-2008-11-12/pdf/E8-25654.pdf).*

- Did the modeling take into account diurnal air traffic patterns? Due to the differences in dispersion characteristics at different times of day discussed above, diurnal patterns may make a significant difference in impacts.

> *EPA Response: Air quality modeling at the model airport did incorporate the diurnal profile in air traffic, as well as hourly meteorology.*

- According to Figure B-1, the ME full LTO fraction at RHV peaks in the late afternoon (4:00 PM on weekdays and 5:00 PM on weekends), but that peak is not seen with SE full LTOs. This distinction may be important because of the approximately 10X greater impacts of ME than SE planes, as shown in Table 1. Was this diurnal difference used in the modeling, or just to select the runway end with the highest total piston engine traffic?

<u>EPA Response</u>: EPA incorporated this diurnal profile difference between ME and SE activity in the air quality modeling (see Appendix A for details on air quality modeling at the model airport).

- Again, AQFs were generated by averaging the 12 values calculated. Since the NAAQS is a not-to -exceed value and the activity for the most active runway for the most active 3-month period is used in the analysis, shouldn't the analysis also use the AQFs that correspond to that period, in order to take into account seasonal variations in meteorological conditions?

  <u>EPA Response</u>: We have responded to this comment above.

4. Section 3.3 presents an analysis to refine estimates of piston-engine aircraft activity using airport-specific data for a subset of airports. Please comment on whether there are alternative airport-specific data, or analysis approaches, that could improve estimates of piston-engine aircraft activity at a subset of airports. In addition, please comment on whether parameters other than piston-engine aircraft activity should be included in analyses to potentially improve model-extrapolated concentrations at a subset of airports, noting that additional parameters are evaluated at all airports in the uncertainty and variability analyses presented in Section 4.

   The refined analyses used more airport-specific factors for fraction of piston planes and for breakdown between SE and ME. I would like to see modeling done at a subset of the airports to look at the impacts of some of the other factors discussed above (e.g. diurnal variation).

   <u>EPA Response</u>: EPA acknowledges that to take steps beyond extrapolated estimates, conducting airport-specific air quality modeling is an approach that would provide a robust assessment of lead concentrations attributable to aircraft emissions. Additional modeling is outside the scope of this effort; however, we have identified the key airport-specific parameters that would need to be incorporated in such modeling if further assessment of specific airports is conducted.

5. EPA provides coarse comparisons of monitored lead concentrations to model-extrapolation results from the national and airport-specific analyses, in Sections 4.1 and 4.2, respectively. In Section 4.3, EPA provides a more detailed comparison of data from lead monitors placed in close proximity to the locations of model-extrapolated concentrations. Please comment on the appropriateness of the approaches to compare model-extrapolated results to monitored concentrations of lead given available monitoring data. Based on your understanding of the methods presented in the report and the comparisons of monitor and model-extrapolated concentrations, please provide your perspective on the performance of the methods in

characterizing the ranges of lead concentrations from piston-engine aircraft at and downwind of US airports.

I'm not clear what point is being made by the discussion in 4.1 about the model showing decreased concentration with distance and with lower levels of piston-engine activity. Wouldn't that always be the case for a ground-level volume source? If the modeling included multiple sources or an elevated stationary source, maximum impacts may occur at some distance from the highest source. However, in this case, only aircraft using the busiest runway end were modeled and both the run-up and take-off operations take place at that end. Similarly, it also seems obvious that the per plane modeled impacts from ME planes would be higher than for SE planes, since the emissions are substantially higher.

> *EPA Response: As the Reviewer points out, the model-extrapolation results conform to expectations regarding pollutant gradients and the relative contribution of SE versus ME emissions rates. We point to these characteristics of the results simply to confirm for readers that the data are in line with these basic measures, particularly for readers who are less familiar aircraft emissions sources and modeling.*

I'm not sure what conclusion to draw from the monitor to model comparison in Figure 7. In most cases, the comparisons look reasonable, but, from the limited data presented, it is impossible to see the shape of the variability of concentration. Would it be possible to model with a receptor at the location of the monitors?

> *EPA Response: EPA understands the interest in seeing the gradient in lead concentrations from extrapolated values at the exact monitoring location. Resource limitations prohibit our ability to run AERMOD at the eight airports shown in these figures (now Figures 9 and 11). The goal in comparing the model-extrapolated lead concentrations with monitored values is to provide information on which the reader can draw a general understanding of the reasonableness of the extrapolated lead concentrations estimates at, and downwind from the area of maximum impact (also, as noted in Section 4.3.2, there are many inherent differences in time, space and methods between these concentrations that need to be taken into account). Also see responses to Comment 6 from Reviewer 2.*

The monitor-model comparability using airport-specific data in Figure 9 appears to be considerably better than the comparability using the national estimates in Figure 7. Again, it would be interesting to model a receptor at the location of the monitor. The fact that the extrapolated modeled results in Panel A of Figure 12 are considerably

lower than, and do not appear to be well correlated with, the monitored values is not reassuring.

<u>EPA Response</u>: We acknowledge the Reviewer's point regarding the relationship between model-extrapolated and monitored concentrations in Panel A of Figure 12; we note in the report that these model-extrapolated lead concentrations were 12% to 52% lower than the primary monitor concentrations. We have conducted extensive analysis to describe the variability and uncertainty in the extrapolated estimates and it is instructive to note that the lead concentrations at the primary monitor all fall within the quantitative uncertainty bounds provided by evaluating variability in run-up time and avgas lead concentration. EPA is providing the data in Figure 12 to illustrate the ability of the extrapolation method to appropriately identify airports with the potential for lead concentrations to be above the level of the lead NAAQS. To that purpose, the figure presents several rolling 3-month average lead concentration values at specific airports where monitored lead concentrations have consistently violated the NAAQS (panel A) and been consistently below the NAAQS (panel B).

6. Section 4.3 presents quantitative and qualitative uncertainty analyses of the model-extrapolated results provided in previous sections. Please provide your perspective on the methods used to conduct these uncertainty and variability analyses, as well as the key parameters EPA included in the analyses (based on previous work discussed in Section 2). Please provide your perspective on the application of this analysis to further characterize the range of lead concentrations attributable to piston aircraft activity at airports nationwide.

Figure 10 is confusing – the caption doesn't match the figure.

<u>EPA Response</u>: This comment has been addressed.

While the methods used to conduct the uncertainty analysis appear to be reasonable, the results are not convincing. In particular, it appears that run-up time durations have such a dramatic effect on impacts that calculations of impacts using the AQFs would not provide useful estimates of impacts at other airports unless they are adjusted to account for differences in that parameter. Note that the document states that "run-up emissions accounted for 82% of the 3-month average lead concentration attributable to piston-engine aircraft in EPA air quality

modeling", that the "variation between the 5th and 95th percentiles of average run-up times observed in EPA modeling resulted in an almost 8-fold variation in concentration attributable to only run-up emissions" and that "the model-extrapolated concentrations from the airport-specific activity analysis are consistently at or near the 2.5th percentile of the Monte Carlo bounds while the 97.5th percentile of the Monte Carlo analysis is up to 2-fold higher than the model-extrapolated concentrations from the airport-specific activity analysis" due to the use of "a much shorter run-up time in developing the model-extrapolated lead concentrations in the national analysis compared with run-up times that have been observed at other airports." Other factors, including wind speed variations, further add to this uncertainty.

Therefore, although the analysis presented in this paper provides information about the range of impacts, I am skeptical about the advisability of using the AQFs for calculating impacts and, potentially, compliance with the NAAQS, at individual airports.

> *EPA Response: As the Reviewer points out, and, in alignment with the goals of this assessment, the report provides information about the range of estimated concentrations of lead in air at and downwind of the maximum impact area at airports nationwide. We further refine theses estimates at a subset of airports where there may be the potential for lead concentrations to be above the level of the lead NAAQS. EPA is not using these estimated lead concentrations to evaluate attainment of the lead NAAQS.*
>
> *We incorporated additional text in the Introduction to the Report to further emphasize that in making determinations regarding violations of the NAAQS for lead, EPA relies solely on the lead surveillance monitoring network. EPA's guidance on this matter is provided in the National Ambient Air Quality Standards for Lead Final Rule (*http://www.gpo.gov/fdsys/pkg/FR-2008-11-12/pdf/E8-25654.pdf*).*

3.4 Response to Comments Received from Reviewer 4: John R. Pehrson

CDM Smith

The general description of piston-engine aircraft operations, use of lead in AvGas, and the lead impact locations at airports is adequate for this evaluation. The discussion clearly states the reason for the study, the general approaches used to develop the lead inventories and air quality factors, and comparison to monitored concentrations at the model airport.

However, an item of potential concern is the fairly consistent model under-prediction of monitored lead concentrations at the maximum impact location and at locations 60 meters downwind (shown in Figure 1 on page 9) at the model airport. The report identifies three likely sources of uncertainty: (i) exact location of run-up activity relative to monitoring station locations, (ii) duration of run-up activity for each flight, and (iii) whether single-engine or multi-engine aircraft were being used. A recent study also highlights the variability of piston-engine aircraft emissions for the same engine type and pilot (Yacovitch et al. 2016):

> "In contrast, piston engines, which drive small propeller planes, operate in a much more fexible manner. Piston engines are rugged and imprecise and pilots can operate them in various ways with simple levers (e.g., the throttle and mixer) in the cockpit. Power and emissions are weakly linked, particularly in low-power states like idle and taxi. The nature of piston engines means that there is also a great deal of variability in their emissions, *even for the same pilot operating the same airplane*." [Chapter 3, page 11]

This is worth noting more clearly in the early sections of the USEPA report. The findings presented in Yacovitch et al. indicate that measurements at the same location, with the same aircraft, pilot, and wind conditions could still produce noticeably different results.

> *<u>EPA Response</u>: We appreciate the information provided in Yacovitch et al., and we now cite these findings as a contribution to variability and uncertainty in Section 4.4. The variability in individual pilot behavior and aircraft emissions that the Reviewer points to could influence concentrations both higher or lower than modeled concentrations. To the extent that the modeling at RHV is under-predicting lead concentrations, we acknowledge the Reviewer's point that this is not necessarily tied only to the three parameters previously highlighted. We have added text to Section 4.3 to highlight additional sources of variability; we also added text to Section 2 to further clarify that the under-prediction in modeling*

results at the model airport is within the common model evaluation criterion of values falling within 2:1 of monitored concentrations. Notably, these are daily values, which inherently include more variability than the 3-month rolling averages used to calculate model-extrapolated concentrations at airports nationwide. Comparisons between model-extrapolated concentrations and monitor data, presented in Sections 4.1, 4.2, and 4.3 also suggest that any under-prediction at the model airport still yields general agreement between model-extrapolated and monitored concentrations.

Returning to other causes of uncertainty, Appendix A discusses sensitivity analyses conducted to determine the potential cause of the modeled discrepancy with monitored results. On pages 22 through 25 of Appendix A, the three likely sources of uncertainty were noted and sensitivity analyses were conducted. Run-up duration is one of the parameters considered in the sensitivity analyses. It appears from the data presented in Appendix A, ten days of activity data including engine run-up durations were collected – all in the Summer season. Would weather conditions encountered in other seasons prompt pilots to spend more time in engine run-up before taking off?

EPA Response: There may be a short increase in run-up time for carbureted engines to conduct de-icing during winter conditions. We account for variation in run-up duration, which could result from differences in seasonal conditions, individual airport characteristics, or other attributes, by conducting a Monte-Carlo analysis that draws from a distribution of run-up times across five airport studies. Further, the maximum activity period for a given airport is likely to be in the summer, as GA operations nationwide peak in May and reach a minimum in January (Wang and Horn. 1985).

Wang, G. H., & Horn, R. J. (1985). Temporal patterns of aircraft operations at US Airports: A statistical analysis. *Transportation Research Part A: General*, *19*(4), 325-335.

Following this discussion, two other sources of modeling uncertainty were noted beginning on page 25 of Appendix A: selection of the initial sigma-y and sigma-z values, and source exclusion zones for volume sources. It is not clear if any sensitivity evaluations were conducted on these parameters. The model exclusion zone issue can be avoided if area sources were used to model engine run-ups instead of volume sources. The size of the area would require some thought,

although selecting parameters that mimic the initial sigma-y value could be used. With regard to choosing appropriate sigma-y and sigma-z values, one study measured those parameters (Wayson, et al. 2003) and presented values for turboprop commuter aircraft as well as large commercial aircraft.

EPA Response: We appreciate the Reviewer's comments regarding the choice of volume vs. area source for aircraft and the selection of initial sigma-y and sigma-z values. These are parameters for which we did not conduct sensitivity analysis. We selected volume sources to treat piston-engine aircraft emissions given the similarity between exhaust from these engines and motor vehicles for which volume sources are typically applied when evaluating the near field environment. This approach is consistent with other piston-engine aircraft emission studies available in the literature (Carr et al. 2011, Heiken et al. 2014, Feinberg et al. 2016). We are unaware of any published work evaluating the treatment of piston-engine aircraft sources as area vs. volume sources. The treatment of aircraft sources as area or volume sources is a modeling topic noted for future research, particularly when evaluating concentrations in the near field environment from jet engine operations (Arunachalam, et al., 2017).

Arunachalm, S., Valencia, A. et al., (2017). Dispersion Modeling Guidance for Airports Addressing Local Air Quality Health Concerns. *National Academy of Sciences Airport Cooperative Research Program*. Research Report 179.

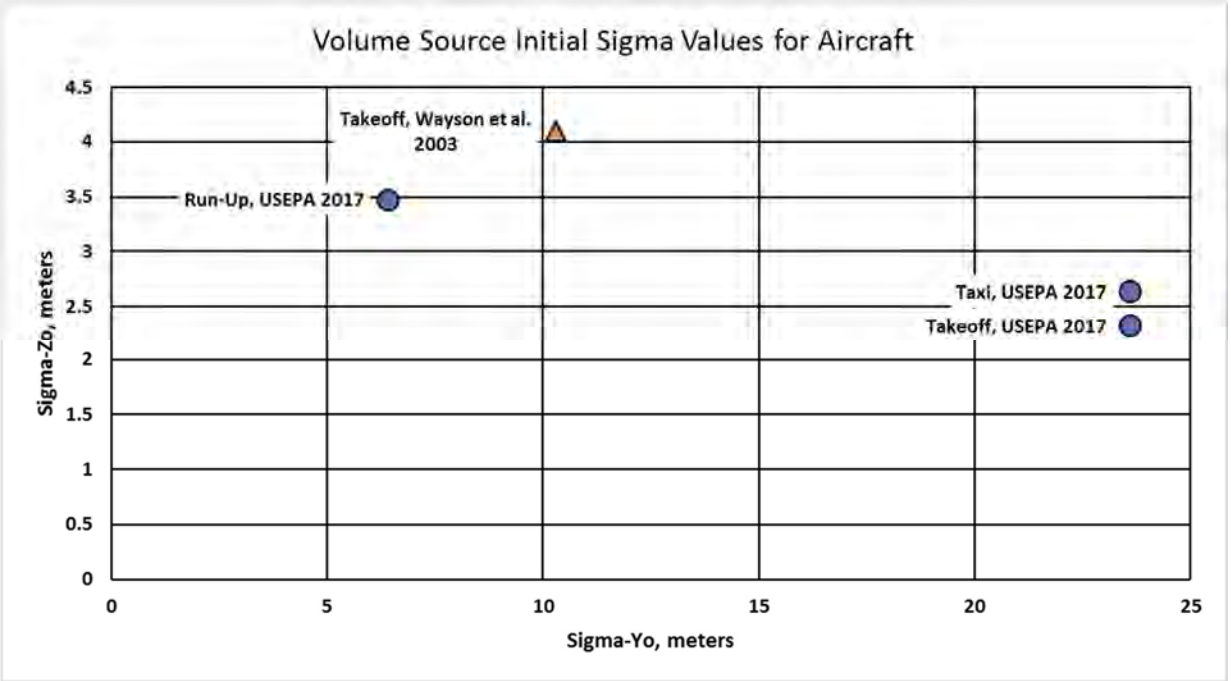
Carr, E., et al. (2011). "Development and evaluation of an air quality modeling approach to assess near-field impacts of lead emissions from piston-engine aircraft operating on leaded aviation gasoline." Atmospheric Environment **45**(32): 5795-5804.

Feinberg, Stephen, and Jay Turner. "Dispersion Modeling of Lead Emissions from Piston Engine Aircraft at General Aviation Facilities." *Transportation Research Record: Journal of the Transportation Research Board* 2325 (2013): 34-42.

Heiken, J., et al. (2014). *Quantifying Aircraft Lead Emissions At Airports.* ACRP Report 133.

http://www.nap.edu/catalog/22142/quantifying-aircraft-lead-emissions-at-airports.

The estimation of initial sigma-y and sigma-z values for volume source representation of aircraft lead emission dispersion, the values could be compared to the values for commuter aircraft (mostly turboprops) developed from a LIDAR study of aircraft exhaust plumes presented in Wayson et al. Typical turboprop initial sigma-y values during takeoff were 10.3 meters, and initial sigma-z values were 4.1 meters during the takeoff roll (from 52 measurement events for commuter aircraft). The measured values from Wayson, et al., 2003 implies that the sigma-y values for takeoff in the report under review may be high by a factor of approximately 2, and the taxi sigma-y value may also be high in the current report. The chart below compares the values from the two reports.



Given that commuter turboprops operating from a large commercial airport are usually multiengine aircraft with seating capacity for 10 to 20 passengers, the sigma values presented in Wayson et al., for takeoff would likely over-estimate initial sigma values for typical general aviation aircraft. The USEPA study under review indicates a substantial difference in takeoff/taxi initial sigma-y values to those for run-up operations. If such a relationship holds across different aircraft sizes, and if the takeoff sigma-y value for turboprops from Wayson et al. is more appropriate for

general aviation aircraft, then the run-up initial sigma values presented in this USEPA report might also be over-estimated. As noted on page 25 of Appendix A, such an overestimation of initial sigma values would likely result in underprediction of concentrations at nearby receptors.

> *EPA Response: We agree with the Reviewer's evaluation that, in comparison to Wayson's measurements of initial sigma values during take-off for turboprops, the initial sigma values we selected for modeling piston aircraft taxi and takeoff might be over-estimated, and could additionally contribute to the underestimate of lead concentrations at our model airport. We note this as a source of uncertainty in Section 4.4.2.*

While the text in the last paragraph on page 8 implies that seven (7) days of modeled and monitored data were compared, Appendix A, page 3 (middle paragraph) states that only three (3) days had both surveyed hourly operational data and lead monitoring data, and the other four (4) days of monitored concentration data used an average activity profile developed from 10 days of activity surveys to predict the monitored concentrations. This may be a point worth noting (footnoting?) in Section 2. In addition, it is not clear which three days of simultaneous measurements and activity data were used since the list of days with collected activity data included four (4) days of simultaneous activity data and lead measurements (8/20, 8/23, 8/26, and 8/28) – per Appendix A, pages 1 and 2, and footnote 1.

> *EPA Response: The three days of simultaneous air monitoring and activity data collection were 8/23, 8/26, and 8/28; southerly winds on 8/20 resulted in operations occurring predominantly on Runway 13L where we did not have monitors placed to evaluate the maximum impact and downwind gradient in lead from piston aircraft; therefore, model-to-monitor comparisons were not used for this day. Clarification on this point was added to Footnote 1 in Appendix A.*

Potential typographical error: Section 1, Page 4, 2nd Paragraph. In the center of the paragraph it is stated that: "Piston-engine aircraft conduct approximately 62 million landing and takeoff operations (LTOs) annually (USEPA 2011)." In reviewing the data referenced (the 2011 NEI data site), it appears that the piston-engine LTOs included in Table 5 of EPA-420-B-13-040 is approximately 32 million LTOs, not 62

million. Not sure if this is a typo, or if the 62 million figure is meant to also include touch-and-go operations and/or helicopter operations.

> *EPA Response: The original sentence mistakenly referred to the 62 million individual operations (landing and take-off events separately) as LTOs. We have corrected the sentence to state that 32 million LTOs were conducted by piston-engine aircraft.*

The description of the approach to develop AQFs was clear. The AQFs were developed from modeled results for each operation type (LTO vs T&G) and aircraft size (SE vs ME). I'm assuming that there isn't sufficient monitoring data to conduct a multivariate regression analysis on daily monitored concentrations to tease out AQFs.

> *EPA Response: This assumption is correct.*

I found Table 2 to be quite understandable. The steps to disaggregate total operations down to daily/hourly operations, assignment of operations to specific runway ends, and aggregating the results for the most commonly used runway was clearly described in Table 2. Given that the detailed hourly data is difficult to obtain, or is non-existent, the approach used in this study is well considered. Figures 2 and 3 also clarified the approach.

The methods used to estimate airport-specific activities and SE to ME ratio appears reasonable.

The findings presented in Figures 7 and 9 are rather compelling, given the range of potential uncertainty in the parameters used to estimate lead concentrations. These figures clearly indicate that the overall approach to estimating lead concentrations at general aviation airports produces results that are fairly consistent (same order of magnitude and often within the expected range) with measured values at a handful of airports with lead monitoring stations.

My primary concern with the results presented in Section 4 is the apparent assumption that the maximum impact site is considered ambient air, especially since the concentrations estimated at a distance of only 50 meters from the maximum impact site already fall below the lead NAAQS. Footnote 40 at the bottom of page 32 provides the official definition of ambient air by USEPA. I believe that USEPA generally accepts that ambient air begins outside of access-controlled areas of an air emissions source, at least for stationary source permitting. If the airport were considered the source, and if a fence is installed around the airport, then

ambient air would occur along the fence line. Given that the airport operator does not generally own the aircraft using the airport, this typical definition of ambient air may not apply. However, if this definition is considered to apply, then most of the "maximum impact sites" for lead concentrations noted in this study would not be ambient air if the general public does not have unrestricted access to these sites.

> *EPA Response: Ambient air is defined by EPA regulations as that portion of the atmosphere, external to buildings, to which the general public has access. At airports, the general public includes recreational pilots and their passengers, members of the public who visit the airport for special events, and may include other populations (e.g., people who rent hangars). Locations at airports to which this population has access include parking lots, observation decks, hangars, and access roads to hangars.*
>
> *For purposes of this report, instead of evaluating individual airports for areas of potential ambient air, we have elected to apply the simple criterion of unrestricted access. The final report identifies only those facilities for which there is unrestricted access within 50 meters of the maximum impact site at airports with model-extrapolated lead concentration estimates above the level of the lead NAAQS.*

The document should provide a clear description of what is meant on Page 6, 2nd paragraph: "*Following EPA practice*, this analysis focuses on the maximum impact area at airports nationwide…" While it is technically easier to calculate lead concentration near aircraft run-up areas, it may not be the appropriate point to estimate impacts relative to ambient air quality standards.

> *EPA Response: We revised this section of the report to be more clear. It is common practice, particularly in a screening analysis, to evaluate the maximum impact locations. And, with regard to the NAAQS for lead, EPA specifically noted the need to evaluate maximum impact locations (see source-oriented monitoring in https://www.gpo.gov/fdsys/pkg/FR-2008-11-12/pdf/E8-25654.pdf).*

Editorial comment: Figures A-1 and A-2 are not very clear. These figures may each need to be full page to see all of the detail discussed in the figure captions.

> *EPA Response: We have increased the size of the figures.*

Potential typo: Appendix A, Page 20, Equation A-4: Is the term "*Eq. 4*" in Equation A-4 really meant to be *"Equation A-3"*?

EPA Response: We have corrected this error in the report.

Format consistency: The text on page 5 of Appendix A uses dots (.) in the table numbers: "… in Table A.1…" while the actual table titles use dashed (-): Table A-1. Might check this throughout the document.

EPA Response: The references to the table have been corrected for consistency.

Potential typo: Appendix C, Page 4, 2nd Paragraph: A figure is suggested, but not actually shown at: "The concentration from only run-up emissions at the maximum impact site receptor was 0.034 µg/m3 for the 5th percentile, 0.257 µg/m3 for the 95th percentile, and 0.092 µg/m3 for the default run-up duration, as shown in *Figure C-*. Lead concentrations…"

EPA Response: We have corrected this typo in the report.

References

Wayson R.L., Fleming, G.G., Kim, B., Eberhard, W.L., Brewer, W.A., Draper, J., Pehrson, J., and Johnson, R., 2003. The Use of LIDAR to Charaterize Aircraft Exhaust Plumes. Air & Waste Management 2003 Annual Conference and Exhibition. Paper No. 69965.

Yacovitch, T.I., Zhenhong, Y., Herndon, S.C., Miake-Lye, R., Liscensky, D., Knighton, W.B., Kenney, M., Schoonard, C., and Pringle, P., 2016. ACRP Report 164 - Exhaust Emissions from In-Use General

3.5 Response to Comments Received from Reviewer 5: Sandy Webb

Environmental Consulting Group LLC

Overall the report was well written with straight forward, systematic, and detailed methodology descriptions. The "way finding" (section introductions, how sections were organized, descriptions of section contents, etc.) throughout the report was very helpful.

1. **Sections 1 and 2 describe the nature of how piston-engine aircraft operate for safety and logistical reasons, along with the previous work that EPA and others have conducted to characterize concentrations of lead in air at individual airports. As stated in the report, conducting detailed air quality modeling or monitoring at all US airports is not feasible due to resource constraints. Please comment on the extent to which this information is clearly described and provide your perspective on the approach selected to utilize modeling from an individual airport in order to characterize concentrations of lead in air at and downwind of maximum impact areas of airports nationwide.**

The methodology described is very systematic, rational, and clearly written. Each section was clearly introduced and the sequence of information was logical.

Page 4, paragraph 4 describes clearly the source and purpose of lead in avgas. However, it is perhaps an over simplification. Tetraethyl lead is "splash blended" into unleaded gasoline resulting in highly variable lead concentration. This can be seen in Appendix C. While the spec is for 2.12 grams/gallon, it can be quite a bit higher and lower. Also this was the only mention of 100 octane avgas. More justification for not modeling the higher lead-content gasoline should be provided.

> *EPA Response: We have added explanation to Appendix C regarding the rationale for not modeling the higher lead-content avgas in this assessment. Appendix C also includes discussion of variability in avgas lead concentrations.*

It would be very useful in Section 1 to include at least a qualitative material balance for lead into and out of a piston engine – describing TEL in gasoline into the engine, combustion transforming the lead, lead captured in engine oil, and lead in exhaust. Include a physical description of the lead emissions (Are there any aerosol emissions containing lead? Is there information on particle size distribution? Are all lead emissions as elemental lead?) A more quantitative explanation of some of this

appears in Appendix A, page A3, Aircraft Emission Rates. This will help explain the importance of fuel consumption rates and times-in-mode.

> *EPA Response: We have included information requested by the Reviewer in Appendix A.*

In Section 1 paragraph 2 and footnote 5 as well as Section 2.1 in the last paragraph on page 7, the run-up area is declared the "maximum impact area." While this is likely true, a more complete explanation is warranted. From these descriptions it was not clear whether ground-level lead concentrations in the maximum impact area were based on modeling, which would include emissions dispersed into the run-up area from other parts of the airport, or if these concentrations were solely the result of run-up emissions. I believe it was the modeled emissions but was confused by the write up in Section 2.1 Appendix A, Table A-1 shows much higher monthly and annual emissions from taxi-out than run-up. Also, taxiways, parallel to runways, will largely be upwind of the run-up area for whichever runway end is in use. Were these emissions captured in the model airport modeling that was the basis for the AQF? Could taxi-out emissions be a significant addition to run-up emissions in the maximum impact area? Whether or not there is a significant contribution this should be explained. It was difficult to tell from reading the AQF methodology (Equation 1) if it was a top-down computation from airport-wide emissions calculations or a bottoms-up computation based on run-up activity. I may well have missed something in the explanation but it was not clear to me.

> *EPA Response: To quantitatively evaluate the relative contribution of aircraft lead emissions during different modes of operation on lead concentrations in air, EPA modeled the emissions, including locations and relative emission rates of lead at the model airport. The taxi-way emissions were included in the development of the AQFs at the model airport. The Reviewer makes an important distinction between the relative contribution of run-up and taxi-out lead emissions versus the contribution of these emissions to concentrations in the defined area of maximum impact near the run-up location. We added clarification to the Introduction of the report to note that while taxi-out emissions of lead can be higher than lead emissions during run-up, the taxi-out emissions occur while the aircraft is underway, which causes much greater dispersion of the emissions compared with the ground-based emissions during run-up which occur while the aircraft is stationary. Thus, the quantitative contribution of run-up emissions to lead concentrations in air at the maximum impact site is much larger than the contribution of*

taxi-out emissions at this location. However, in order to holistically evaluate aircraft emissions on air quality, EPA has included all aircraft emissions when developing AQFs from the study at the model airport (i.e., taxi and other modes are included in AQFs along with run-up) (we added a Footnote in Section 3.1 to clarify this point).

Section 2.2 says "…the model tended to under predict monitored concentrations …" Systematically under predicting monitored concentrations raises questions about the methodology. See also Figure 8. This was practically dismissed as a concern. Perhaps a more complete explanation is warranted. I was unable to discern the reason.

<u>EPA Response</u>: EPA has added additional text to the description and discussion of the model-to-monitor comparisons conducted at the model airport. See response to Reviewer 3 on Page 32. We note that this underestimation may be due to day-to-day operational variability, changes in the location of run-up procedures, and variability in the monitored concentration, among other reasons addressed in the response to Reviewer 3; however, this under prediction does not appear to be systematic considering modeling performed at other airports (see references in Section 2.2) using the same or similar approaches. Additional discussion has been added to Sections 2.2 and 4.4 to address sources of uncertainty and their implication on the analysis.

The Reviewer also points to the airport-specific analysis later in the report with the reference to Figure 8 (now Figure 10). There, we expanded our discussion of the evaluation of individual airports in Section 3.3. This expanded discussion now includes a series of sensitivity analyses to capture potential sources of uncertainty and variability. We examine airports where concentrations may be underestimated in the national analysis by re-evaluating these airports with the assumption that all GA activity and 50% of AT activity would be performed by piston-engine aircraft. This sensitivity analysis, in effect, increases the maximum modeled concentration by a minimum of 28% for each airport. The seven days of model-to-monitor comparison at the monitor airport showed a model underestimation of, on average, of 16.8%. Thus, we believe our sensitivity analyses are, to first order, sufficient for capturing uncertainty from potential model underestimation at the model airport.

Regardless of these comments, I do think the overall methodology is a good one for calculating lead emissions for GA airports nationwide.

2. **Section 3.1 presents the methods to calculate Air Quality Factors (AQFs) at the model airport. Please comment on the approach used to calculate AQFs at the model airport specifically for the purposes of using these factors to estimate concentrations at and downwind of maximum impact areas of airports nationwide.**

Developing the AQFs by aircraft class and applying those to other airports is a legitimate approach and should provide consistent results from airport to airport. However, as noted above I couldn't tell whether the AQFs were computed based on total airport emissions or only those from the run-up area. If the AQFs are appropriately developed this procedure for estimating concentrations makes sense and works well for a national analysis.

> *EPA Response: We appreciate the Reviewer's perspectives on the approach and confirm here that all aircraft emissions (i.e., during all modes of operation) were included in developing the AQFs.*

3. **Table 2 and accompanying text in Section 3.2 describe the methods used to estimate piston- engine aircraft landing and take-offs (LTOs) at individual runway ends on a rolling 3-month basis (e.g., apportioning out piston-engine-specific LTOs from total LTOs at each airport, allocating annual activity to daily and then hourly periods). Are these methods clearly described and do you have recommended changes to the steps taken? Please explain any alternative options and provide the location of data sources that would support such alternative options.**

Table 1 in Section 3.1 would be easier to read if the data was presented as (10^{-6} mg Pb/m^3/LTO).

> *EPA Response: This comment has been addressed.*

Table 2 in Section 3.2 was very systematic and clear. Including the rationale for each step was very helpful. Overall the table is so detailed and long it is difficult to keep all steps in mind. The calculations for number of operations by aircraft type by hour, month, and averaging period all seemed reasonable and appropriate for developing a national analysis. This process appears to get the proper result so no suggestions for alternative options or data sources.

I found Step 13avi in Table 2 confusing. Does this imply the need for and availability of avgas lead concentration at each airport or is this only for the model airport?

> *EPA Response: Step 13avi in Table 2 describes the step introduced to provide extrapolated lead concentrations at airports where lead concentration in avgas has not been measured. The avgas lead concentration at the model airport was used in the air quality modeling results that provided the AQFs, which were in turn used to calculate model-extrapolated concentrations at each airport nationwide. In order to use a standardized lead avgas concentration instead of the concentration measured at the model airport, we scaled model-extrapolated concentrations at each airport by the ratio of the ASTM maximum standard to the concentration at the model airport. The result provides concentrations of lead in air at each US airport that include an avgas lead concentration equal to the ASTM maximum.*

4. **Section 3.3 presents an analysis to refine estimates of piston-engine aircraft activity using airport-specific data for the subset of airports. Please comment on whether there are alternative airport-specific data, or analysis approaches, that could improve estimates of piston-engine aircraft activity at a subset of airports. In addition, please comment on whether parameters other than piston-engine aircraft activity should be included in analyses to potentially improve model-extrapolated concentrations at a subset of airports, noting that additional parameters are evaluated at all airports in the uncertainty and variability analyses presented in Section 4.**

National data sources of data on GA activity are notoriously unreliable because of the nature of much of the GA activity. There is significant training activity at many GA airports but often no compelling reason to track T&Gs. Many aircraft are flown infrequently. Data reports to FAA frequently are not prepared rigorously. All emissions analysis projects regarding general aviation have to deal with these limitations. The most accurate analyses are based on data collected on-site. This is not feasible for preparing a national estimate of lead concentrations. As far as my experience with GA data sources, EPA has selected the most reliable and the uncertainty analysis suggests that data limitations will not have a significant effect on the overall estimation results.

On page 29, the discussion of Figure A-9 references concentrations "…labeled as Heiken Fuel Consumption…" There are no references to Heiken in the table.

<u>EPA Response</u>: This comment has been address and the reference has been corrected.

Figure A-9 Run-Up and the discussion in paragraph 2 on page 31 – Based on this data, the run-up fuel consumption (time-in-mode and consumption rate) used in this analysis looks low compared to Heiken et al's findings. Also for the discussion of run-up times-in-mode in Section C.2, especially the last sentence in paragraph 1 on page 4 in C.2.2, "Their modeling found that changing the emissions attributable to run-up from 3% of modeled emissions to 5% of modeled emissions resulted in a 34% increase in annual atmospheric lead concentrations." Perhaps the run-up fuel consumption should be reconsidered and the concentrations recomputed using a higher TIM and consumption rate. This could account for the lower modeled concentrations compared to the measured concentrations.

<u>EPA Response</u>: The time in mode (TIM) data at the model airport are based on observations at that facility, and while we agree that the fuel consumption rate for run-up emissions used in the air quality modeling at the model airport is low compared with data from Heiken et al., any increase in fuel consumption during the run-up mode of operation would improve the model-monitor comparisons for some of the days of under-prediction, but would result in an over-prediction for some days. Given the sparse dataset for fuel consumption during run-up (i.e., only four unique fuel consumption rates across engines types compared with 18 unique fuel consumption rates for other modes of operation), we did not include this parameter in a quantitative uncertainty analysis. As noted by a separate Reviewer, fuel consumption rates and times in mode vary from pilot to pilot and within events for the same pilot. We note parameters other than TIM and fuel consumption rates that may contribute to the model-to-monitor differences and quantify their impacts in Appendix A. When extrapolating the model airport results to other airports, we account for differences in TIM during run-up by conducting a quantitative uncertainty analysis to characterize the impact of this important parameter on lead concentrations in maximum impact areas of airports nationwide. While a sufficient sample size was available to conduct such a quantitative analysis for run-up TIM, data were insufficient to conduct a similar analysis on fuel consumption rate.

The second paragraph on page 4 in Section C.2.2 references a figure with no figure number ("Figure C-"). It appears there is a missing figure showing sensitivity analysis.

EPA Response: This comment has been addressed.

5. **EPA provides coarse comparisons of monitored lead concentrations to model-extrapolation results from the national and airport t-specific analyses, in Sections 4.1 and 4.2, respectively. In Section 4.3, EPA provides a more detailed comparison of data from lead monitors placed in close proximity to the locations of model-extrapolated concentrations. Please comment on the appropriateness of the approaches to compare model-extrapolated results to monitored concentrations of lead given available monitoring data. Based on your understanding of the methods presented in the report and the comparisons of monitor and model-extrapolated concentrations, please provide your perspective on the performance of the methods in characterizing the ranges of lead concentrations from piston-engine aircraft at and downwind of US airports.**

The model-extrapolated approach is probably the best approach for estimating lead concentrations at airports nation-wide. The modeled vs. monitored and model airport analysis confirm this. As noted in Section 4.2, only about 27 out of 13,000 airports are likely to exceed the NAAQS. These are all high activity airports and will likely have the best data on operations, based aircraft, mix of aircraft type, etc.

As noted in an earlier comment, having the model-extrapolated estimates falling consistently below airport-specific analysis raises a concern, however, the reason for this is not evident.

EPA Response: We believe the Reviewer is raising two potentially inter-related points. The first, in referring to a previous comment, notes the potential systematic under-estimation in modeled concentrations at the model airport. As discussed in the response to the comment about Section 2.2 above, modeling conducted by EPA using similar methods indicates that the approach used does not systematically underestimate monitor concentrations. We have added text to that section and Section 4.4 to better indicate that we recognize model performance as a source of uncertainty and its implications on interpreting these results. This comment may also pertain to the data we provide regarding the uncertainty analysis around the national analysis results which indicate that the model-extrapolated concentration estimates are consistently lower that alternative values that use the range of data available for run-

up duration. This is because the run-up duration at the model airport was at the low end of the range of values measured at other airports.

The second point refers more specifically to the evaluation of individual airports for potential to have lead levels above the lead NAAQS. In the airport-specific analysis, we compare two model-extrapolated lead concentrations for each of the airports with the potential to have lead levels above the lead NAAQS. The national analysis estimates typically (but not for all airports) fall below the airport-specific estimates as a result of replacing national average piston-engine activity fractions with airport-specific data. A potential explanation for this is that many airports with a significant amount of GA activity may have a higher percentage of GA activity performed by piston-engine aircraft than at other airports. This occurrence would result in the model-extrapolated concentrations in the airport-specific analysis exceeding the estimate for the model-extrapolated concentrations using national default piston-engine percentages. Thus, the airports that are identified in the airport-specific analysis may represent airports that both have high activity and a high percentage of piston operations. We recognize that activity and operational data, both using national default and airport specific values, is a source of potential uncertainty, and we have added text to better describe activity uncertainty to Section 4.4 of the report.

Page 47, last paragraph, second sentence says "As a conservative assumption, the ASTM standard for the maximum lead concentration in 100LL was used in the national analysis …" As shown in Figure C-1, the lead concentration is often higher than the standard so I would not consider this a "conservative" assumption, however, I think it is a reasonable assumption.

EPA Response: We agree and have deleted the use of this word in the report in reference to concentration estimates.

6. **Section 4.3 presents quantitative and qualitative uncertainty analyses of the model-extrapolated results provided in previous sections. Please provide your perspective on the methods used to conduct these uncertainty and variability analyses, as well as the key parameters EPA included in the analyses (based on previous work discussed in Section 2). Please provide your perspective on the application of this analysis to further characterize the range of lead concentrations attributable to piston aircraft activity at airports nationwide.**

The analysis methods used to assess the analytical results are appropriate and seem to be well carried out. My experience and knowledge of statistical techniques is rather limited and I will rely on the other reviewers to assess this section. I know some of them are experts with these techniques.

The description of the results of Monte Carlo analysis gets a bit tedious and could perhaps be moved to an appendix.

> *EPA Response: We endeavored to keep the text describing the Monte Carlo analysis as streamlined as possible while allowing the reader to understand the basic steps undertaken.*

The caption for Figure 10 is not very clear, especially the explanation of the black and blue concentration values.

> *EPA Response: This comment has been addressed and the figure caption has been improved for clarity.*

The text should point out the different scales used in Figures 10 and 11.

> *EPA Response: This comment has been addressed.*

It is very difficult to read the concentration scales on Figure 12, both panel A and B.

> *EPA Response: We have made changes to the text to make it easier to read.*

Section 4.3.2 – the first half of the first paragraph is a very clear explanation of the approach taken with this analysis. In the second half of the paragraph, I find the discussion of mixing height confusing. It was not clear to me that mixing height would have much of an effect on concentrations in the run-up area/maximum impact location. In the 3rd paragraph, it would be helpful to give range and not just the difference for the 3-month AQF.

> *EPA Response: We have rewritten Section 4.4.1 to respond to reviewer comments on meteorological parameters and other sources of uncertainty. We have placed the discussion of mixing height and its contribution to concentration uncertainty in the context of other meteorological parameters. We agree with the Reviewer that, given that the largest contributor to concentrations at the maximum impact location is nearby run-up emissions, mixing height would not be the*

biggest contributor to uncertainty. However, we do note that even near-source concentrations of primary pollutants have been shown to be dependent on meteorological variables and these variables may contribute to greater uncertainty at downwind locations.

Section 5, paragraph 2, sentence 1 – provide the max value and put "greater than NAASQ" into the parenthetical statement.

EPA Response: This comment has been addressed.

7. Editorial Comments

Page 4, paragraph 3, delete unneeded words, "… aviation gasoline ~~for several reasons~~, namely to help…"

Page 4, footnote 1 – second sentence should read "Facility types other ***than*** airports …"

Page 8, paragraph 2 – beginning of second line has inconsistent use of a dash "aircraft- and meteorological data," – should be a dash to follow meteorological or not one after aircraft.

Page 13, paragraph after numbered list – line 3 should read "… were used ***to*** evaluate …"

Page 18, line 10 should read "… runway ***is*** multiplied …" rather than "… runway as multiplied …"

Page 20, Table 2, Step 1, step description – delete the first word "Determine"

Page 23, Table 2, footnote 29, last line – the wording after "airports" is very awkward where it says airports "likely have a distinct activity profile from GA airports." This should be explained more clearly/simply.

Page 48, first paragraph, Table 5 on page 50 is too far from the initial table callout on page 48, paragraph 1, line 9

Page 51, first paragraph, figure 10 on page 52 is too far from the initial figure callout on page 51, paragraph 1, line 1

Page 56, paragraph 3, line 6 should read "… order to capture …"

Page 62, references, for the Heiken reference, it should read ACRP Report 133 rather than ACRP 02-34. Once published, ACRP drops all references to project numbers in deference to the report number. This report is also referenced on pages A32, Appendix B references, and C9

Page B1 through end of Appendix B; page numbers are missing

Figure B-1 is split across two pages (B3 and B4) making it look like a figure number is missing on page B3

Figure C-6 (a), (b), (c), and (d) – add Run-Up Time units of measure to y-axis

EPA Response: The editorial comments above have been addressed.