Review of Manuscript

In-situ and Denuder Based Measurements of Elemental and Reactive Gaseous Mercury with Analysis by Laser-Induced Fluorescence. Results from the Reno Atmospheric Mercury Intercomparison Experiment (acp-2016-446)

This manuscript describes the work of Hynes et al., to quantify elemental gaseous mercury (Hg^0) and total gaseous mercury (TGM) concentrations using a sequential two photon laser-induced fluorescence (2P-LIF) instrument off a manifold as part of the RAMIX method inter-comparison study conducted in Reno, NV. As the authors point out, there is currently a debate in the literature concerning the efficacy of various ambient mercury measurement methods under different conditions of ambient relevance. As such the RAMIX study endeavored to provide a platform for a definitive methods comparison. Unfortunately, based on Hynes et al work (and references within) the RAMIX study and the manifold delivery system designed and implemented for this study fell short of this goal in several important aspects that limits the utility of the study's findings. As a result, I believe this paper is an important contribution to the state-of-science. I also have some technical concerns/questions with the implementation of some of the experiments described in the paper enumerated in the comments below, therefore I feel this manuscript will require substantive revision before it is acceptable for publication in ACP.

General Comments:

(1) In several areas of the paper (e.g., lines 63-64; 144-147) the authors incorrectly include <u>particulate bound</u> mercury (PBM) into their definition of total <u>gaseous</u> mercury (TGM). In the literature TGM is generally used to describe Hg^0 + divalent reactive gaseous mercury (RGM). In fact most ambient instruments that quantify TGM (e.g., Tekran Instruments Corporation Model 2537) use an integrated Teflon filter to exclude all particulate matter, and most monitoring networks include an integrated Teflon filter at the inlet of their sampling line to minimize gas/particle interactions in the sampling line that has been shown to create problems in reliably quantifying ambient gaseous mercury species (see below). If the system is pyrolyzing all ambient mercury species to Hg^0 for detection, then perhaps coin a new operationally defined term such as total atmospheric mercury (TAM) to avoid confusion.

(2) The authors need to provide additional details on their sampling configuration and calibration procedures (QA/QC) in the methods section to provide readers and reviewers the basic information necessary to inform and judge the implementation of the their 2P-LIF system.

(3) Though not a problem with this paper per se, the design and implementation of the RAMIX study manifold system described by Finley et al. (2013) for this work is problematic and the results described in this paper must be viewed through this lens. The overarching issue was the decision to allow ambient particulate matter (PM_1) into the manifold. When the stated objective was to evaluate the efficacy of different measurement methodologies to quantify Hg^0 and RGM the logical first step would have been to evaluate these gaseous species first without the complication of spiked gases interacting with ambient particles in the manifold. The curious choice of constructing the manifold using highly porous PFA Teflon tubing also creates multiple potential problems (i) absorption/desorption of spiked gas concentrations following diffusion gradients, (ii) non-conducting material combined with high (187 LPM) flow rate and low relative humidity could lead to electrostatic collection of PM on the internal manifold walls providing additional surfaces for gas phase adsorption, and (iii) poor conduction of heat applied to the external surface of the tubing at such a high manifold flow rate (Finley et al. reported using eight thermocouples to measure external temperature down the length of the manifold but did not report any measuring internal surface or air temperature). Using a blower to maintain flow through the manifold also added a reported 15% uncertainty in the spike concentrations (Finley et al., 2013). All these issues lead to relatively low reported average transmission efficiencies for Hg^0 (92%), $HgBr_2$ (76%), and O_3 (93%) even under controlled laboratory conditions (Finley et al., 2013). The range for Hg^0 transmission in the manifold was reported to be 71-101% by Prestbo (2014). If the spiked mass of target gases into the manifold were not conserved through the system, then definitive evaluation of the analytical instruments sampling from the manifold were compromised. I agree with the authors that 15% probably underestimates the overall uncertainty for Hg^0 (Lines 205-207) and that the system should be characterized as "a semi-quantitative delivery system" (Lines 210-211). If the manifold cannot quantitatively and reproducibly transmit a relatively inert gas like Hg^0 there is very little chance of reproducibly transmitting $HgBr_2$.

(4) This reviewer agrees with the authors that previous RAMIX study papers invoke and discuss mercury oxidation and reduction chemistry that is not supported in either the theoretical or applied literature. Controlling the physical adsorption of gaseous species onto manifold surfaces and aerosols is the logical first direction for which to find answers.

Specific Comments:

(1) Introduction:
   a. Lines 49-50: Sprovieri et al., 2016 Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-466 is a more contemporary discussion of global background Hg^0 concentrations in the Northern and Southern hemispheres from the global GMOS network.
   b. Lines 63-64: PBM is not part of TGM.

(2) Experimental:

a. As indicated above, the authors need to supplement the QA/QC information.
    i. Lines 122-126: The authors discuss calibrating the 2P-LIF system using a Tekran 2537B as a secondary transfer standard, and that unlike their previous field work the unit was turned off for one week for transportation to the study site. This discussion implies to the reader that there was some kind of additional uncertainty in the calibration due to the Tekran 2537B unit being powered down. As long as the authors powered up the unit and had argon purge gas flowing for 2-3 days prior to use the permeation system should have re-equilibrated and had no impact on their 2P-LIF system calibration. I suggest the authors clarify the circumstances of the Tekran 2537B operation status. Manual standard addition injections from the Tekran Model 2505 primary calibration source (that the authors indicated they had on site – Line 180) should have been conducted to verify the stability/accuracy of the 2537B instrument perm tube system prior to initiation of the experiment. The authors should report results for any QA/QC injections.
    ii. Lines 122-126: Clarify if the 25' sampling line was heated and shielded from the sun.
    iii. Lines 130-131: If the Tekran 2537B was not able to pull 2 LPM with the additional load of a 25' sampling line it is not clear why the authors simply did not reduce the flow rate set point to 1.5 LPM (the manufacturer recommended flow rate). Adding a supplemental pump to the instrument exhaust and increasing the instruments internal vacuum can impact the permeation tube system performance creating uncertainty in the instruments reported values. The additional vacuum may also have played a role in the permeation tube "malfunction" described by the authors in Lines 230-231. Was the Tekran 2537B checked with manual standard addition injections from the Tekran Model 2505 primary calibration source in the external pump configuration?
    iv. Lines 133-134: The authors indicate that the 2P-LIF system cannot detect RGM, but expressed concern about "deposition of RGM on the sampling lines followed by heterogeneous reduction to GEM". Were any actions taken to filter out RGM from the sampling line (at the manifold port) while allowing the GEM to pass to the instrument like incorporation of a soda and lime trap? Allowing $HgBr_2$ into a Tekran 2537B can have long term contamination effects on the internal components (filter packs, tubing, connectors, and valves).
    v. Lines 148-150: Was a second pump used to maintain flow through the sampling line not being actively sampled by the 2P-LIF system during the TOM difference experiment? Otherwise air in the sampling line void volume not being actively sampled would stagnate and not represent the correct temporal sample duration.

vi. Lines 152-169: How often were the KCl-coated manual annular denuder cleaned and recoated/conditioned?
vii. Lines 231-234: The authors discuss comparison of the UM and UNR Tekran 2537 instruments and bring up the point again about the power down of their 2537B instrument. The authors should clarify how long the system was allowed to stabilize prior to the first perm tube calibration, and how often thereafter it was recalibrated. Was an independent QA auditor part of the study plan to ensure traceability across research group instruments?

b. Line 146: Again PBM is not part of TGM.
c. Line 157: 2003 should be 2002.
d. Line 165: Not heating the KCl-coated annular denuders during sampling can be problematic since the hydrophilic KCl coating will tend to absorb water vapor and can (i) interfere with RGM collection, and (ii) provide surfaces for heterogeneous reactions.

(3) Results:

a. Lines 231-236: In the absence of an independent auditor or standard addition injections to validate the respective instruments perm tube emission rates, it is not possible to definitively establish the reason for the observed differences between the UM and UNR instruments. Based on the described behavior it could be related to contamination of one or both of the instruments with $HgBr_2$ or an unstable permeation tube system. It would be useful for the authors to discuss the observed behavior as a function of the timing of $HgBr_2$ spiking.
b. Lines 300-301: The authors point out in this discussion that the UM 2537B was systematically reporting a higher Hg^0 value than the UNR instrument. If the authors believe the divergence between 2537 instruments was due to the UM instrument being turned off for shipment, then the UM instrument would be reporting lower values. This would be due to the fact that the amount of mercury being emitted by the unequilibrated perm tube system would be higher than expected during the calibration since excess Hg^0 accumulated on the walls of the perm tube oven would be slowly desorbing – resulting in lower reported ambient concentrations.
c. Line 399: The authors are discussing Fig. 7 in this discussion and then state “In addition, it is clear that the DOHGS system show very different temporal profiles of TOM.” I suspect the authors should reference Fig. 9 here since I do not see the DOHGS concentrations presented in Fig. 7.
d. Line 413: I do not see the DOHGS concentrations presented in SI Figure 6.
e. Lines 422-423: I do not see the DOHGS concentrations presented in SI Figures 7-9 as referenced in the text.
f. Lines 447-455: It is unclear what the authors take home message for Section 3.5.3 discussion. It is also unclear why the Spec2 data are shown

in SI Fig. 13 since it was sampling off the manifold while all the other measurements are from the trailer roof.

g. Lines 489-530: Landis et al. 2002 documents the quantitative transport of $HgCl_2$ through the manual denuder elutriator/impactor inlet when properly heated.
h. Lines 490-492: The authors implementation of the manual denuders method described by Landis et al., 2002 deviated in two ways (i) not using the suggested elutriator/impactor inlet to remove large aerosols which may be retained by the denuder causing positive artifacts (the potential contamination of their RGM denuder sample by PBM is later discussed by the authors), and (ii) not heating the denuder system.
i. Lines 493-494: Feng et al., 2004 reference cited by the authors to imply potential loss of RGM by the inlet elutriator/impactor inlet of the manual denuder system does not support their statement. The Feng et al. paper does not use or even mention this system or RGM loss in their particulate mercury methods paper. They used a method described by Lu et al., 1998 that does not use an inlet.
j. Lines 517-522: Manual versus automated denuder and denuder in series experiments described here were previously conducted and presented in Landis et al., 2002. No significance difference between manual and automated systems, and no significant breakthrough from the first denuder. This previous work should be cited.
k. Lines 579-582: The critical review of the experimental design here is warranted and should go further to include recommendations to (i) improve the manifold design and sampling port configurations as previously discussed, (ii) include an independent auditor, and (iii) removal of PM from the gas phase experiments.