We have placed the comments from the referees in italics. Our replies are in normal type. New text is in normal type unless it is an addition to a sentence and here the additions are in bold type.

Referee One

"1) As mentioned in the paper, the LIF signal must be calibrated to convert the fluorescence signal to an absolute concentration. The data in this paper appear to have been calibrated by reference to a commercial CVAFS instrument rather than using a calibration standard. In addition, the signal is shown to be sensitive to the concentration of water vapor. Unfortunately there is little discussion regarding the absolute calibration of the instrument, although it appears that an expanded discussion will appear elsewhere (pages 5661 and 5668). Given the importance of accurate calibrations, the authors should include an expanded discussion on how they plan to calibrate the instrument in the laboratory and in the field."

We discussed calibration in "Section 3.3: Calibration" and "5.3 Permeation Oven Calibrations". We noted that our Tekran is equipped with the internal calibration option which is based on a gravimetrically calibrated permeation tube. We demonstrated that standard addition to the sample flow is possible but using an uncalibrated permeation tube because we were unable to obtain a gravimetrically calibrated Hg(0) permeation tube prior to RAMIX

We have expanded "Section 3.3: Calibration" adding the following:

"In addition standard addition of Hg(0) to the sample flow could be introduced by rerouting the sample flow through the calibration permeation oven. We were unable to obtain a gravimetrically calibrated Hg(0) permeation tube prior to RAMIX and, as a result, this did not offer an independent calibration. However the Tekran and 2P-LIF responses were compared as described below. For future instrument deployments we will utilize a low output, gravimetrically calibrated permeation tube to offer an independent absolute calibration of the Hg(0) concentration that will offer an additional check on the Tekran response. "

1

"However, the stability of the laser system appears to be an issue that is currently limiting the overall precision and accuracy of the measurements."

2) Much of the variability in the measurements appears to be the result of instabilities associated with the laser wavelength and power (page 5660).

We did not mean to imply that fluctuations in the measurements are a result of instabilities. The later portion of Section 5.2 has a discussion of this. We write: "Fig. 10 shows the average of the two signals and the reference cell output normalized to the same amplitude as PMTs A and B. It is apparent that there are fluctuations on a 7 second timescale that are captured by both LIF PMTs but some of the fluctuations are not correlated. Comparison with the reference cell output shows that the variation in the ambient LIF signal is not due to pulse-to-pulse or short term wavelength variations in the probe lasers. In fact the mean of the normalized signals is 1.7 ng m⁻³ and the 1 σ variation in the signals is 0.18, 0.17 and 0.03 ng m⁻³ for PMT A, PMT B and the reference PMT, respectively."

To avoid giving the impression that instability is a problem we have modified "3.2 Reference Normalization" adding the sentence in BOLD:

"3.2 Reference Normalization

At the start of a measurement cycle the excitation dye lasers were optimized to produce the maximum 2P-LIF signal by maximizing output power and then tuning the wavelengths to generate the maximum 2P-LIF signal. At this point the laser wavelengths should be centered on the Hg(0) absorption features and the 2P-LIF signal will be proportional to the Hg(0) concentration. In practice both the wavelength and power of the excitation lasers can drift and the 2P-LIF signal will no longer be proportional to the Hg(0) concentration. **Typically after an initial startup period we found the laser systems to be remarkably stable for several hours without requiring adjustment. Nevertheless** the power will decrease as the laser dyes degrade and wavelengths can drift as a result of temperature variation and problems with mechanical stability."

We have also modified the discussion in section 5.2 adding:

"The comparison with the reference cell output shown in Fig. 10 demonstrates that the variation in the ambient LIF signal is not due to pulse-to-pulse or short term wavelength variations in the probe lasers. There is almost no variation in the reference cell 2P-LIF signal during this period."

"To account for these changes, the authors normalize the measurements to changes in the reference cell signal. However, this normalization depends on the stability of the concentration of Hg(0) in the reference cell. It also assumes that any saturation of the fluorescence signal is similar in the reference cell compared to the sampling cell. Did the authors measure the Hg(0) from the reference cell using one of the Tekran instruments to test its stability? Have the authors demonstrated that changes in the reference cell signal with respect to laser power and wavelength were similar to changes in the sampling cell?"

We monitored the temperature of the reference permeation oven but not the stability of the output and we acknowledge that it would be useful to check this. We verified that the 2P-LIF reference cell tracked changes in the sample cell when the crystal alignment or wavelength were changed but an additional systematic study of these effects would be useful. We have added text:

"We found that changes in crystal alignment and wavelength shifts reducing the 2P-LIF signal by a factor of five were corrected by normalizing to the reference cell. However we have not performed a systematic investigation of the relative responses of the sample and reference systems to changes in power and wavelength drift and we plan to do this as part of any future deployment. "

"What impact would multiple pulses on the same airmass in the reference and sampling cell have on the fluorescence signal in terms of saturation of the transition? Would this impact the ambient (roof) measurements (where the flow of air is variable)?" I'm not completely sure what the concern is here. A significant fraction of the Hg atoms will be excited during a single 5 nsec laser pulse but the lifetimes of all these states are very short so all excited atoms will have returned to the ground state before the next laser pulse even at the higher 50Hz rep rate.

"Minor comments:

1) Given the complexity of the instrument, the paper would benefit from a schematic figure illustrating the various components of both the "first generation" system and the "second generation" system."

We have added a block diagram of the "first generation" system. The only significant changes in the second generation system are a change in the pump and output wavelength of one dye laser and we think it adds to confusion to attempt to incorporate this in the diagram

"2) It would be useful to provide more details regarding the lasers, PMTs, etc. used in the instrument (manufacturer, model numbers, etc)."

This has been done

"3) What laser powers were used for each wavelength, and how did they vary? Was there a significant drop in laser power between the reference cell, the detection cell, and the roof measurements? Information regarding the dye concentration and solvent used should be provided."

This has been done

"4) More details regarding the geometry of the sampling cell and reference cell should be provided, perhaps with a figure. What was the flow rate through the sample and reference cells?"

We have not added a figure. We added:

"The sample and reference cells were identical and constructed of 1" i.d. pyrex. They were 12" long with two 1" side arms attached at the midpoint. O-ring joints (#25) at the ends of both the main cell axis and the side arms allowed windows to be attached and easily removed for cleaning or exchange."

The flow rate through the sample cell (10 SLPM) was in the original manuscript, the flow through the reference cell (100 cc/min) has been added.

"5) The configuration of the roof-top measurements is not clear. On page 5658 it states that two PMTs were used for roof sampling, while on page 5670 it appears only one PMT was used for the ambient measurements. This should be clarified in the text and in the schematic figure."

One PMT was used for sampling at RAMIX, whilst two were used in testing the "second generation" system in Miami. Text has been added to clarify:

"During RAMIX a single PMT was located on the trailer roof. For sampling with the "second generation" system in Miami two PMTs were located on the trailer roof. After exiting the roof the beam then passed ~1 cm in front of the photocathode of the PMT or PMT's that were located ~ 3ft above the trailer roof. "

Referee Two

Given that the first goal of this paper is the description of the experimental set-up, I am missing details such as laser energy, beam sizes (for example you say one beam has been expanded for better overlap with the second beam: is the second dye laser producing a larger beam?) and PM voltage (do you change the voltage applied to the PM regularly and why? This is not clear, especially not how this is taken into account in the calibration). A schema of the set-up would be helpful.

A schematic is now included, and more detail is provided on the experimental setup including names of instruments, laser power and beam size.

The manuscript already contains the following text:

"The voltage and hence the gain of the PMTs was adjusted to optimize the sensitivity and dynamic range of the 2P-LIF signal however tube B had higher gain and required more frequent adjustment. All data was processed to account for the varying amplification of the PMTs, then background subtracted."

This simply means that we established gain curves, i.e. PMT sensitivity as a function of voltage for each PMT and then normalized the signals to account for any change in PMT voltage to put all signals on the same scale.

"Being not an Hg-expert, I had trouble following the different excitation schemes (page 5655 and 5656): maybe a figure showing the different transitions would help to make it easier?"

A figure has been included.

"At this point I would appreciate an explanation about why you use two different excitation schemes? I was guessing that generating the 2nd generation laser wavelengths is easier (and you say so at the end, page 5672), but it would be good to have an explanation here. Maybe there are other reasons?"

We are simply trying to improve the sensitivity to get better sensitivity so that with the 50 Hz system we can average 10 shots and attain the precision required for eddy-correlation measurements. We have added:

The first generation system was deployed at RAMIX. In an effort to investigate the possibility of increasing sensitivity for high frequency measurements we have made preliminary measurements using a "second generation" system.

"Throughout the manuscript I was wondering why you use two PM, and little by little I understood that this is for better signal. Please explain this when you mention the 2 PM for the first time (Page 5658). You mention that the set-up could be improved by using even more PM: do you have plans to do so? How about using a lens to collect more photons?"

We have added the following text:

"The use of two PMTs allowed the dynamic range of the detection system to be increased, increased the number of photons detected and, as discussed below, allowed us to distinguish between real fluctuations in Hg(0) concentration and random fluctuations associated with photon statistics." We are currently seeking funding to do additional deployments of the instrument. The PMTs are rather expensive but we hope to be able to purchase more. Even with a fluorescence cell the PMTs are close to the excitation beam because there are no filters. We have tried lenses but get no improvement perhaps because an increase in collection efficiency is offset by transmission losses. Even S1-UV fused silica has losses at 185 nm.

Same page, you use a cold trap (what temperature): is this just for condensing water? Are you sure no Hg can be trapped?

We have added:

"If the ambient temperature is significantly above the trailer temperature and the humidity is high the sample air is passed through a cold trap, **set in an ice-bath**, to prevent condensation on the sampling lines inside the trailer."

Mercury is almost insoluble in water (this is the reason that elemental mercury is not scavenged by precipitation and has a very low deposition rate). The vapor pressure of mercury at 200 K is higher than its atmospheric partial pressure so no mercury will be trapped at 273 K.

Same page, you use 254nm instead of 253.7 and 408 instead of 407.8: please use same wavelength everywhere.

This has been done.

Page 5659: you introduce the RAMIX manifold: the short explanation on this manifold that you give page 5666, would be better placed here.

We have added:

"The manifold is described in detail by Finley et al. [2013]."

Page 5663/4: you discuss the signal in air compared to helium. The paragraph is not very clear to me. You say before that the quenching rate for the higher states is not known, so what values did you use to calculate the expected fluorescence efficiency of 0.5%?

We have added text:

As noted above in the low laser power linear fluorescence regime we would expect a fluorescence efficiency of less than 0.5% based on quenching of the $6^{3}P_{1}$ state, additional quenching of the $7^{1}S_{0}$ and $6^{1}P_{1}$ states would reduce the fluorescence efficiency even more.

In "Section 2.0" we note:

The $6^{3}P_{1}$ state has a radiative lifetime of 119 nsec and the quenching rate coefficient with O_{2} is 3.6 x 10^{-10} cm³ moecule⁻¹ s⁻¹ [Michael and Suess, 1974]; Breckenridge and Umemoto, 2007]. The fluorescence efficiency in air at atmospheric pressure can be calculated using the Stern-Volmer relationship and gives a fluorescence efficiency of 4.7×10^{-3} and an effective radiative lifetime of 0.56 nsec.

Abbreviations are sometimes not well defined: Please define RGM the first time you use it (line 10 in the abstract). Line 21 of the abstract you could already introduce the abbreviation TGM.

This has been done.

What is a KCl denuder (page 5654, line 9)?

The text has been changed to:

The current approach to measurement of RGM relies almost exclusively on sampling on KCl-coated annular denuders followed by pyrolysis and CVAFS analysis of the Hg(0) produced by RGM decomposition [Landis et al., 2002].

The Landis et al. paper describes the technique in great detail and includes detailed schematics of the types of annular denuder that are used.