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.

Anthony J. Hynes[*], Stephanie Everhart, Dieter Bauer, James Remeika, and Cheryl Tatum Ernest[1]

Division of Marine and Atmospheric Chemistry, Rosenstiel School of Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker Causeway, Miami, Florida 33149.

[1]current address:Atmospheric Chemistry Department, Max Planck Institute for Chemistry, Hahn-Meitner-Weg 1, Mainz, 55128, Germany

[*] Corresponding author. Tel.: +1-305-421-4173; fax: +1-305-421-4689. E-mail address: ahynes@rsmas.miami.edu (A. J. Hynes)

Abstract

The University of Miami (UM) deployed a sequential two photon laser-induced fluorescence (2P-LIF) instrument for the in-situ measurement of gaseous elemental mercury, Hg(0), during the Reno Atmospheric Mercury Intercomparison Experiment (RAMIX) campaign. A number of extended sampling experiments, typically lasting 6-8 hours but on one occasion extending to ~24 hours, were conducted allowing the 2P-LIF measurements of Hg(0) concentrations to be compared with two independently operated instruments using gold amalgamation sampling coupled with Cold Vapor Atomic Fluorescence Spectroscopic (CVAFS) analysis. At the highest temporal resolution, ~5 minute samples, the three instruments measured concentrations that agreed to within 10-25%. Measurements of total mercury (TM) were made by using pyrolysis to convert total oxidized mercury (TOM) to Hg(0). TOM was then obtained by difference. Variability in the ambient Hg(0) concentration limited our sensitivity for measurement of ambient TOM using this approach. In addition, manually sampled KCl coated annular denuders were deployed and analyzed using thermal dissociation coupled with single photon LIF detection of Hg(0). The TOM measurements obtained were normally consistent with KCl denuder measurements obtained with two Tekran speciation systems and with the manual KCl denuder measurements but with very large uncertainty. They were typically lower than measurements reported by the University of Washington (UW) Detector for Oxidized Hg Species (DOHGS) system. The ability of the 2P-LIF pyrolysis system to measure TM was demonstrated during one of the manifold $HgBr_2$ spikes but the results did not agree well with those reported by the DOHGS system. The limitations of the RAMIX experiment and potential improvements that should be implemented in any future mercury instrument intercomparison are discussed. We suggest that instrumental artifacts make a substantial contribution to the discrepancies in the reported measurements

35 over the course of the RAMIX campaign. This suggests that caution should be used in drawing significant

36 implications for the atmospheric cycling of mercury from the RAMIX results.

38

40 **1.0 Introduction:**

41   The environmental and health impacts of mercury pollution are well recognized with impacts on

42 human health and broader environmental concerns (U.S. EPA., 2000; UNEP, 2013; Mergler et al. 2007;

43 Diez, 2009; Scheuhammer et al., 2007). There have been extensive reviews of global emissions,

44 measurements and biogeochemical cycling of mercury, (Mason, 2009; Streets et al., 2011; Pirrone et al.

45 2009; Lindberg et al., 2007; Ebinghaus et al., 2009; Sprovieri et al., 2010; Selin, 2009) The concerns

46 associated with the mercury problem have resulted in attempts to regulate and control emissions at both

47 national and international levels. The latest attempt in the United States is incorporated in the Mercury and

48 Air Toxics Standards (Houyoux, and Strum, 2011; US EPA, 2013) and international efforts by the United

49 Nations Environment Program have led to the Minamata Convention on Mercury, a global

50 legally binding treaty on mercury controls (UNEP, 2008; UNEP, 2013; UNEP, 2014).

51 There is a reasonable consensus on typical background concentrations of atmospheric mercury, which are

52 extremely low. Typical concentrations range from 1.2–1.4 ng m^{-3} in the Northern Hemisphere and 0.9–1.2

53 ng m^{-3} in the Southern Hemisphere and appear to be decreasing (Slemr et al., 2011; Sprovieri et al., 2016) [

54 1 ng m^{-3} is ~ 3×10^6 atoms cm^{-3} or ~ 120 ppq (parts per quadrillion)]. Until recently it has been accepted

55 that most of the mercury found in the boundary layer is elemental mercury, Hg(0) (Lindberg et al., 2007).

56 Oxidized or reactive gaseous mercury (RGM), normally assumed to be in the Hg(II) oxidation state, has not

57 been chemically identified and is thought to constitute a very small fraction of the total mercury

58 concentration although recent work (Gustin et al., 2013; Ambrose et al., 2013) challenges this view. Our

59 overall understanding of the atmospheric chemistry of mercury and the detailed elementary chemical

60 reactions that oxidize Hg(0) is poor (Lin et al., 2006; Hynes et al., 2009; Subir et al., 2012) and the

61 uncertainty of both the chemical identity and measurements of speciated oxidized mercury places few

62 constraints on models. Atmospheric measurements of mercury represent a significant challenge in ultra-

63 trace analytical chemistry and the issues associated with current techniques have been discussed by Gustin

64 and Jaffe (2010). We have developed a laser-based sensor for the detection of Hg(0) using sequential two-

65 photon laser-induced fluorescence (2P-LIF) (Bauer et al., 2002; Bauer et al. 2014). The instrument is

66 capable of fast, in-situ, measurement of Hg(0) at ambient levels. By incorporating pyrolysis to convert

67 RGM and particulate mercury to Hg(0) it is possible to measure total mercury (TM, i.e the sum of Hg(0)

68 plus gas phase and particulate bound oxidized mercury) and hence to measure total oxidized mercury

69 (TOM, i.e. the sum of gas phase and particulate bound oxidized mercury) by difference. The Reno

70 Atmospheric Mercury Inter-comparison Experiment (RAMIX) offered an opportunity to deploy the 2P-LIF

71 instrument as part of an informal field intercomparison at the University of Nevada Agricultural

72 Experiment Station (Gustin et al., 2013; Ambrose et al., 2013; Finley et al., 2013). RAMIX was an attempt

73 to inter-compare new Hg measurement systems with two Tekran 2537/1130/1135 systems. This is the

74 instrumentation that is currently in use for the overwhelming majority of atmospheric Hg measurements.

75 Participants included the University of Washington (UW), University of Houston (UH), Desert Research

76 Institute (DRI), University of Nevada Reno (UNR) and the University of Miami (UM). The specific goals

77 for the project were:

78   1- Compare ambient measurements of gaseous elemental mercury, Hg(0), gaseous oxidized mercury

79   (RGM) and particulate bound mercury (PBM) by multiple groups for 4 weeks.

80   2- Examine the response of all systems to spikes of Hg(0) and $HgBr_2$.

81   3- Examine the response of all systems to Hg(0) in the presence of the potentially interfering

82   compounds ozone and water vapor.

83   4- Analyze the data to quantify the level of agreement and the results of interference and calibration

84   tests for each measurement system.

85 In practice the instrument operated by UH only measured Hg(0) for the first week of the campaign and the

86 cavity ring down spectroscopy (CRDS) instrument deployed by DRI did not produce any data. Hence

87 RAMIX was primarily an intercomparison of the UM 2P-LIF instrument, the UW Detector for Oxidized

88 Hg Species (DOHGS) that is based on two Tekran 2537 instruments, and a Tekran 2537 and two

89 2537/1130/1135 speciation systems deployed by UNR. Under these circumstances we were not able to

90 compare 2P-LIF measurements made at high temporal resolution with the CRDS instrument. It did allow us

91 to compare the 2P-LIF sensor with independently operated instruments that use preconcentration on gold

92 coupled with analysis by CVAFS and to examine potential interference effects. Our focus here is to

93 compare the short term variation in GEM on the timescale that the CVAFS instruments operate, ~ 5 minute

94 samples, and examine the ability of the different instruments to capture this variation. In addition, we made

95 measurements of TM and hence TOM by difference and also employed manual denuder measurements to

96 attempt to measure RGM directly. In prior publications, Gustin et al. (2013) and Ambose et al. (2013)

97 provide their interpretation of the RAMIX results and their conclusions have very significant implications

98 for our understanding of atmospheric mercury chemistry. In this work we offer a contrasting view with

99 different conclusions.

100 **2.0 Experimental**

101 **2.1 RAMIX Intercomparison.** A detailed description of the RAMIX location and the local meteorology

102 was provided by Gustin et al. (2013). The original RAMIX proposal included participation from Tekran

103 Corporation to build and test a field-deployed, high-flow sampling manifold that could be reliably spiked

104 with 10-100 parts per quadrillion of RGM. Tekran proposed to supply both RGM and Hg(0) spiking using

105 independent generators that were traceable to NIST standards and would be independent of the detection

106 systems being evaluated. However, due to time constraints Tekran believed that it was unlikely that the

107 manifold and ultra-trace spiking system could be manufactured and fully tested to their standards, so they

108 declined to participate in RAMIX (Prestbo, 2016).   Instead, the UW group stepped in to supply and

109 operate the sampling manifold and spiking system and the details of its characterization are provided in

110   Finley et al. (2013). During the RAMIX campaign the 2P-LIF instrument sampled on 18 days, typically

111   sampling for between 4 and 6 hours. The longest period of continuous sampling lasted for 26 hours and

112   occurred on September 1st and 2nd. Over this 18 day period we sampled from the RAMIX manifold and, in

113   addition, at the end of the campaign we sampled ambient air independently and also attempted to measure

114   TOM by pyrolyzing the sample air and measuring the difference between Hg(0) and TM. We also sampled

115   RGM using KCl coated annular denuders using LIF for real-time analysis.

116   **2.2 The 2P-LIF system**

117   Bauer et al. (2002, 2003, 2014) provide a description of the operating principles of the 2P-LIF

118   instrument. Bauer et al. (2014) provide a detailed description of the 2P-LIF instrument deployed at RAMIX

119   including the sampling configurations, data processing, calibration and linearity tests together with

120   examples of experimental data. In summary, the system uses sequential two-photon excitation of two

121   atomic transitions in Hg(0) followed by detection of blue shifted LIF. The instrumental configuration at

122   RAMIX utilized an initial excitation of the Hg 6^3P_1-6^1S_0 transition at 253.7 nm, followed by excitation to

123   the 7^1S_0 level via the 7^1S_0 - 6^3P_1 transition at 407.8 nm. Both radiative decay and collisional energy transfer

124   produce population in the 6^1P_1 level. Blue shifted fluorescence was then observed on the strong 6^1P_1 - 6^1S_0

125   transition at 184.9 nm using a solar blind photomultiplier tube (PMT). By using a solar blind tube that is

126   insensitive to laser scatter at the excitation wavelengths very high sensitivity is possible. The use of

127   narrowband excitation of two atomic transitions followed by detection of laser-induced fluorescence at a

128   third wavelength precludes the detection of any species other than Hg(0). The 2P-LIF instrument requires

129   calibration, so Hg(0) was also measured with a Tekran 2537B using its internal permeation source as an

130   absolute calibration. We sampled from the RAMIX manifold, which was below ambient pressure, through

131   ~25 ft of ¼ in Teflon tubing. No filter was placed on the sampling line to attempt to remove ambient RGM

132   or the $HgBr_2$ spikes that were periodically added to the sample flow. The sampling line was not heated and

133   was not shielded from the sun. The original RAMIX plan called for all instruments to be located close to

134   the manifold for optimal sampling. Unfortunately the positioning of the trailers at the actual site precluded

135   this and forced us to use a long sampling line. As a result, the internal pump on our Tekran was not able to

136   draw the 1.5 SLPM required for sampling and an auxiliary pump was placed on the Tekran exhaust to boost

137   the flow. Under atmospheric conditions the 2P-LIF instrument cannot detect RGM so, in principle, this

138   does not need to be removed from the sample gas. However, deposition of RGM on the sampling lines

139   followed by heterogeneous reduction to GEM could produce measurement artifacts. The limit of detection

140   for Hg(0) during RAMIX was ~30 pg m^{-3} for a 10 s or 100 shot average.

141

142   **2.3 Measurements of TM and TOM**

143   We attempted to use the 2P-LIF instrument to measure TM and hence TOM by difference. Although we

144   have routinely used this approach to convert $HgCl_2$ and $HgBr_2$ to Hg(0) in the laboratory, this was our first

145   attempt to measure total oxidized mercury at ambient concentrations. A second sampling line was attached

146   to the RAMIX manifold and a pyrolyzer was located directly at the manifold sampling port. The pyrolyzer

147 consisted of an ~0.6 cm o.d. quartz tube, 15 cm in length and partially filled with quartz wool. Wrapped

148 Nichrome wire encompassed an 8 cm section of tube that was heated until the quartz began to glow. The

149 high temperature inside the pyrolyzer reduces both RGM and particulate mercury in the manifold air to

150 Hg(0), which is then monitored by 2P- LIF and gives the sum of oxidized (both gaseous and particulate)

151 and elemental mercury, i.e. TM. Directly sampling from the manifold and measuring ambient Hg(0) then

152 allows the concentration of TOM to be calculated as the difference between the two signals. Both lines

153 were continuously sampled at 10 L/min and the flow to the fluorescence cell was switched between the

154 pyrolyzed and unpyrolyzed sample lines in, typically, 5 min intervals to attempt to track fluctuations in

155 [Hg(0)] that would obscure the relatively small signal increase attributable to TOM.

156 **2.4 Manual Denuder Sampling of RGM**

157 We conducted manual denuder sampling on seven afternoons during the RAMIX campaign to

158 attempt to quantify total RGM, We sampled using both KCl coated annular denuders and uncoated tubular

159 denuders that were then analyzed using programmable thermal dissociation (Ernest et al., 2013). In both

160 cases we monitored the Hg(0) that evolved during RGM decomposition, in real time using single photon

161 LIF. Only the annular denuder results are presented here. The use of denuder sampling coupled with

162 thermal dissociation has been described by Landis et al.(2002) and is used in the Tekran Model 1130

163 Mercury Speciation Units deployed during RAMIX. Air is pulled through a KCl coated annular denuder

164 which captures RGM but transmits elemental and particulate mercury. After a period of sampling, typically

165 one hour, the denuder is flushed with zero grade air and the denuder is heated to 500°C. The RGM is

166 thermally decomposed producing elemental mercury that desorbs from the denuder surface and is then

167 captured and analyzed by a Tekran 2537. The KCl coated annular denuders used here were manufactured

168 by URG Corporation and were identical to those described by Landis et al for manual sampling. They were

169 located on top of one of the RAMIX instrument trailers a few feet from the entrance to the RAMIX

170 manifold inlet. The denuders sampled at 10 SLPM, they were not heated and the integrated

171 elutriator/acceleration jet and impactor/coupler described by Landis et al. and incorporated in the Model

172 1100 speciation unit were not placed on the denuder inlet. Hence no type of particle filtering was used on

173 the inlets. The denuders were cleaned and recoated prior to the RAMIX deployment. Prior to sampling, the

174 denuders were cleaned by heating to 500 °C and then bagged and taken to the sampling site. After a period

175 of sampling that varied from ~1 to 4 hours, the denuders were capped, placed in sealed plastic bags, and

176 transported to the analysis lab at the University of Nevada, Reno. On most of the sampling days a single

177 denuder was opened and then immediately bagged serving as a field blank. On the final two days of

178 sampling, denuders were sampled in pairs, i.e with two denuders connected inline so that the front denuder

179 sampled RGM and the rear denuder served as a blank and monitor of bleed-through of RGM. The blank

180 concentrations are typically low as shown in Table 1, however on September 10th the blank shows a very

181 high value that is indicative of significant contamination at some point during the cleaning or sampling

182 process. For the analysis, a flow of He passed through the denuders and then into a fluorescence cell where

183 any Hg(0) in the flow was detected by LIF. The LIF was monitored by two PMT's set to different gains to

184 increase the dynamic range of the detection system. Prior to the analysis, a known amount of mercury was

185 injected into the flow through a septum using a transfer syringe. The syringe sampled from a Tekran Model

186 2505 Mercury Vapor Primary Calibration Unit. Without disrupting the gas flow the denuder was then

187 placed in a clamshell tube furnace that had been preheated to 500°C. The evolution of the Hg(0) was

188 monitored for, typically, 5-10 minutes and after the LIF signal had returned to baseline a second calibration

189 injection was performed. A frequency doubled, Nd-Yag pumped dye laser was used to excite the Hg(0)

190 6^3P_1-6^1S_0 transition at 253.7 nm and resonance LIF was observed at the same wavelength. In this approach,

191 the detection PMT detects both LIF and laser scatter, hence sensitivity is limited by the ratio of intensity of

192 the LIF signal to the laser scatter. Since the 6^3P_1 level is efficiently quenched by both O_2 and N_2

193 (Breckenridge and Unemoto, 2007) the thermal analysis was performed in He buffer gas to achieve good

194 detection sensitivity. The excitation beam then passed through a reference cell that contained a steady flow

195 of Hg(0) from a permeation source. The LIF signal from the reference cell served to confirm that the laser

196 output was stable.

197 3.0 **Results:**

198 3.1 RAMIX Manifold

199 As noted above, the RAMIX manifold had to be constructed and tested by the UW group under tight time

200 constraints and details of its characterization are provided in Finley et al. (2013).. A critique of the

201 manifold performance has been presented by Prestbo (2014) and we detail some key issues here. The

202 manifold deployed at RAMIX was a different size than the prototype tested in the laboratory. The

203 laboratory manifold showed very large variation in calculated transmission efficiencies of Hg(0) after

204 spiking with a permeation source. Finley et al. reported recoveries of 71-101% for short-term spikes. The

205 authors speculate that this was associated with rapid changes in ambient Hg(0) but provide no

206 measurements to support this. The Hg(0) source used for spiking was gravimetrically calibrated by the

207 manufacturer but was not used at the calibration temperature requiring the output to be calibrated by a

208 Tekran 2537B. After the equipment was moved to the RAMIX site the permeation tube output increased.

209 The authors also acknowledge a significant uncertainty (± 15%) in the RAMIX manifold flow

210 measurements that were required to calculate spike concentrations; hence this is the minimum uncertainty

211 in calculated spike concentrations.

212 In fact, we find that several independent measurements of Hg(0) spikes differ by as much as 30%

213 from the value calculated by the manifold operators suggesting that (± 15%) underestimates the

214 uncertainty. Because of these considerations we believe the RAMIX manifold is best treated as a semi-

215 quantitative delivery system that was not well characterized. We do not feel it is appropriate to

216 characterize "recoveries" as Gustin et al. (2013) have done because of the large uncertainty in Hg(0) spike

217 concentrations. Rather, it is most useful to focus on sampling periods when multiple independent

218 instruments show reasonable agreement.

219

220 **3.2 UM Tekran Performance**

221   In evaluating the first week of the UM RAMIX measurements it became clear that there was some non-
222   linearity in the relative responses of the 2P-LIF and UM Tekran systems and that better agreement was
223   obtained by referencing the Hg(0) concentration to the UNR Tekran. Gustin et al., (2013) concluded that
224   the UNR Tekran, based on the inlet configuration, only measured Hg(0) and they suggested that the UM
225   system, due to the long sampling line, was measuring total gaseous mercury (TGM). We compared the
226   manifold Hg(0) readings from the UM and UNR Tekrans over the first 260 hours in which we took
227   measurements. The absolute concentration difference relative to the UNR instrument is shown in Figure 1.
228   Hour zero corresponds to 9 am on August 26th when we started measurements and hour 260 corresponds to
229   midnight on September 5th. Over the first 24 hours the UM Tekran is offset by ~0.5 ng m^{-3} and the offset
230   jumps to ~ 2 ng m^{-3} at hour 30 on August 27th with the difference decreasing over the next week of
231   measurements in an almost linear fashion. Over most of this period the UW Tekran did not report Hg(0)
232   measurements other than a small set of measurements on August 28th that are offset by ~0.5 ng m^{-3} relative
233   to the UNR Tekran. It can be seen that by hour 250 on September 5th all three instruments had converged.
234   After this period the agreement between the UW, UNR and UM Tekrans was good until September 8th,
235   when the UM instrument became contaminated after a malfunction of our external permeation oven,
236   requiring replacement with a backup Tekran 2537A unit. Both the absolute response and the response
237   factor, i.e. the calibration factor of the UM Tekran were somewhat unstable during this period and
238   additional details are provided in the Supplementary Information. Our focus during this initial period of the
239   intercomparison was on the two laser systems that were being set up. In retrospect we can acknowledge
240   that greater attention should have been paid to quality assurance with the UM Tekran. We conclude that the
241   difference between the UM and UNR instruments is an experimental artifact. Problems with instability in
242   the UM Tekran may have been associated with the use of an external pump to supplement the internal
243   Tekran pump, or with the fact that the UM instrument had been powered down for almost one week and
244   relocated to a site at a significantly different ambient pressure. It is also noteworthy that the initial abrupt
245   change to a large offset followed by the offsets shown in Fig. 1 occurred prior to the start of the manifold
246   spikes of HgBr$_2$ and cannot be associated with the elevated levels of HgBr$_2$ that were introduced into the
247   manifold on Sept. 5th. The differences between the instruments cannot, in our view, be indicative of any
248   type of chemistry within our sampling lines, nor can it be indicative of the UM instrument measuring TGM
249   rather than Hg(0).
250
251   **3.3 2P-LIF Measurements**
252   The absolute Hg(0) concentrations reported for the 2P-LIF measurements typically use a single 10-minute
253   section of Tekran concentration data to calibrate the 2P-LIF signal and place it on an absolute concentration
254   scale. The complete time series of measurements then gives a long-term comparison of the 2P-LIF and
255   Tekran instrumentation with the absolute 2P-LIF concentrations based on the single 10-minute calibration
256   point.
257

3.3.1 September 5th

This was the first occasion on which the three independent Tekran 2537 instruments and the 2P-LIF system reported simultaneous measurements. The 2P-LIF system sampled from the RAMIX manifold for approximately 6.5 hours from ~10:30 am to 5 pm. Over the course of the sampling period there were two spikes of Hg(0) lasting one and two hours, respectively. The UW manifold team reported an initial 10 am Hg(0) spike concentration of 26.5 ng m^{-3} dropping to 24.4 ng m^{-3} over the course of the one hour spike. The two hour spike that began at 1 pm was reported to be ~12.4 ng m^{-3} dropping to 10.5 ng m^{-3} over the course of two hours. The ambient airflow in the manifold was spiked with HgBr$_2$ for the whole of this sampling period and the reported level of the HgBr$_2$ spike varied between 0.6-0.7 ng m^{-3}. The levels of HgBr$_2$ measured by the DOHGS instrument were consistent with this but the concentrations reported by the UNR speciation units were considerably lower and with a significant discrepancy between the two speciation units. Figure 2a shows the sequence of Hg(0) measurements from the UNR, UW and UM Tekrans together with the 5 minute averages of the 2P-LIF signal. The 2P-LIF instrument began manifold measurements in the middle of the initial 10 am Hg(0) spike and is scaled to the concentration at this time which all three Tekrans measured as ~22.5 ng m^{-3}. The three Tekrans agree to better than 5% during both of the manifold spikes and, based on a pre-spike ambient concentration of 2 ng m^{-3} it suggests that the initial spike concentration was ~ 20.5 ng m^{-3}. This suggests that the reported spike concentration was ~25-30% larger than the actual concentration introduced into the manifold. Fig. 2b shows an expanded concentration scale to highlight the nominally ambient measurements. There is some suggestion that it took some time for the spike to be completely removed, particularly after the second spike. At the completion of the second spike all the instruments drop to ambient but the UNR instrument sees two Hg(0) "pulses". Interestingly these show up with greatly reduced amplitudes in the UW and UM Tekran signals and also in the 2P-LIF signal. Figure 3 shows the % difference of the other instruments relative to the UM Tekran and over most of the sampling period the agreement between all the measurements is better than 10% over an ~ 7 hour period with 5 minute sampling resolution. This indicates that the 2P-LIF instrument is capable of stable operation over an extended time period with any drifts being corrected by normalization to the reference cell. Well calibrated independently operated Tekrans should be capable of agreement to better than 5% based on tests performed by the manufacturer and this level of agreement is achieved during subsets of the sampling period. It is not clear if the deviations that are observed, particularly the large deviations seen by the UNR Tekran after the second spike are related to presence of elevated levels of HgBr$_2$, or other issues related to manifold operation. The fact that all the instruments observed these Hg(0) pulses suggests that the artifact may be related to a process in the manifold rather than in in the UNR sampling line. However the significant differences in the magnitude of Hg(0) pulses observed by the different instruments are difficult to rationalize.

3.3.2 September 1st and 2nd

294   The UM and UNR systems sampled simultaneously for a 22 hour period offering an opportunity to
295   compare the instruments over an extended sampling period. This sampling also occurred prior to any of the
296   manifold spikes that introduced substantial concentrations of $HgBr_2$ into the manifold and sampling lines.
297   Unfortunately, the UW instrument did not report any measurements during this sampling period. The UM
298   system sampled for 26 hours and the complete dataset is described elsewhere, (Bauer et al. 2014). This
299   includes a detailed analysis of the short-term, i.e. 1-10 seconds, variation in the Hg(0) concentration and the
300   ability of the 2P-LIF system to capture this. Here we focus on the simultaneous sampling period and the
301   variability that should be resolvable by both of the Tekrans and the 2P-LIF instruments. SI Figure 1 shows
302   the 24 hour sampling period with the 2P-LIF signal calibrated by the UM Tekran concentration at the
303   beginning of hour 13 (i.e 1 pm on September 1st) and the corresponding measurements from the UNR
304   Tekran. SI Figure 2 shows the same data with an expanded y-axis to highlight the variation in the ambient
305   measurements. All three instruments track each other quite well over the first 10 hours and then measure a
306   nocturnal increase in Hg(0) which shows greater medium term variability in the concentration. The 2P-LIF
307   concentrations are approximately 20% greater than the Tekran measurements during this period. At hour 33
308   (i.e. 9 am on September 2nd) there was a manifold spike with a reported concentration of 12.9 ng m^{-3}
309   dropping to 11.9 ng m^{-3} over the course of one hour. The UNR Tekran is ~6% lower, the UM Tekran is
310   ~20% lower and the 2P-LIF ~22% higher than the calculated spike concentration. SI Figure 3 shows the
311   same measurement set but with all instruments normalized to the second manifold spike at hour 33. Figure
312   4 shows an expanded y-axis, the concentration scale, focusing on the ambient concentration measurements.
313   It is apparent that we now see better agreement between the 2P-LIF and the UNR Tekran but that the UM
314   Tekran lies systematically higher than the UNR Tekran. Figure 5 shows a three hour subset of the
315   measurements corresponding to 5-8 am on the morning of September 2nd. The variation between the
316   instruments is greater than 5% and the short term variations in the Hg(0) concentration vary between the
317   three instruments. Using either calibration approach we see that all instruments capture both the nocturnal
318   increase in Hg(0) concentration and the greater variability in the signal but that there are differences in the
319   amplitude of the variability.
320
321   **3.3.3 Hg(0) Intercomparison Conclusions**
322   Almost all of the measurements of atmospheric concentrations of Hg(0) have been made with
323   CVAFS instrumentation and the majority of those measurements have utilized the Tekran 2537. This work
324   provides the first extensive comparison of the Tekran 2537 with an instrument that is capable of fast in-situ
325   detection of Hg(0) using a completely different measurement technique. Measurements over two extended
326   sampling periods show substantial agreement between the 2P-LIF and Tekran measurements and suggest
327   that all the instruments are primarily measuring the same species. Intercomparison precision of better than
328   25% was achievable over an extended sampling period and precision of better than 10% was achieved for
329   subsets of the sampling period. As we discuss below it is difficult to determine the extent to which
330   interferences from RGM contribute to the differences observed.

331

3.4 Interference Tests.

As noted above, one component of the initial RAMIX proposal was an examination of the response of the various sensors to potential interfering compounds $HgBr_2$, O_3 and H_2O. An analysis of the 2P-LIF detection approach suggests that, at the spike levels employed during the RAMIX campaign, neither $HgBr_2$ nor O_3 should have any interference effects. Changes in the concentration of H_2O do affect the 2P-LIF signal because H_2O absorbs the 2P-LIF fluorescence signal and may quench the fluorescence. In addition, O_2 also absorbs the 2P-LIF signal and quenches fluorescence thus a change in the O_2 concentration will affect the linearity of the response. We have presented a detailed discussion of these effects (Bauer et al., 2014) including an examination of two types of interferences that have been observed in LIF sensors applied in atmospheric and combustion environments and concluded that these are not potential problems in 2P-LIF measurements of atmospheric Hg(0). As we have noted previously (Bauer et al., 2014), condensation in our sampling lines can produce artifacts in Hg(0) concentration measurements. Because of the low humidity in Reno it was not necessary to use any type of cold trap during ambient measurements but we did use a trap during manifold spikes of H_2O so our measurements do not address this as a potential interference.

3.4.1 O_3 Interference Tests.

On September 7th an ozone interference test was conducted by simultaneously spiking the sampling manifold with a high concentrations of Hg(0) and ozone. The spike in Hg(0) lasted from 9am to 7:30 pm and there were two ozone spikes, each of two hours duration. A comparison of the UM, UW and UNR Tekrans and the 2P-LIF signal is shown in Figure 6. The UW Tekran only measured for a portion of this period but agrees reasonably well with the other Tekrans. The 2P-LF signal is calibrated by the UM Tekran reading during the initial Hg(0) spike at hour 9.30. The 2P-LIF signal was online for 6 minutes at the beginning of the first ozone spike and then went offline for ~40 minutes for instrument adjustments. When the 2P-LIF came back online the magnitude of the normalized signal was low relative to the Tekrans. At hour 13 all three instruments converge and agree well over the course of the second spike. The magnitude of the 2P-LIF signal could have been affected adversely by the adjustments but any reduction in signal should have been compensated by a corresponding change in the reference cell. The elevated levels of ozone were introduced into the manifold by UV irradiation of O_2 and adding the O_2/O_3 gas mixture directly into the manifold produced a reported ~8% relative increase of O_2 levels in the manifold mixing ratio. As we note above this additional O_2 would absorb some of the 2P-LIF signal but this would be a very small effect. The enhanced quenching by O_2 is more difficult to assess but cannot explain the discrepancy between the Tekrans and the 2P-LIF signal. In addition the agreement during the second ozone spike was good. One possible explanation is that the increase in the O_2 mixing ratio was larger than calculated for the first spike. A second series of O_3 spikes were conducted on September 13th when we were attempting to measure total mercury using pyrolysis as described below. The 2P-LIF measurements switched on a five-minute cycle between a pyrolyzed line that would have decomposed all the ozone in the sample and a line

368   containing the ambient air spiked with ozone. There was no difference in the 2P-LIF signal from the two

369   sampling channels again suggesting that O_3 has no interference effects.

370   The changes in the Hg(0) concentration measurements shown in Figure 6 track the predicted

371   changes in calculated spike concentration. However the calculated spike concentrations, which are also

372   shown are 20-40% higher than the actual measurements made by the Tekrans.

373

374   3.5 Measurements of TM and TOM

375   We made attempts to use the 2P-LIF instrument to measure TM and hence TOM by difference by sampling

376   through two manifold lines. A pyrolyzer was located at the manifold on one of the sampling lines to

377   measure TM. The other sampling line measured ambient Hg(0). TOM was calculated from the difference in

378   the TM and Hg(0) concentrations and in this sampling configuration the limit of detection for TOM

379   depends on the short term variability in ambient Hg(0) which is significant and shows a diurnal variation.

380   The pyrolysis system was set up and tested on September 12. Manifold sampling was conducted on the 13th

381   and 14th and sampling from the trailer roof occurred on the 15th. We calculated the means of the pyrolysis

382   and ambient channel concentrations, and the difference which gives the TOM concentration. We also

383   calculated the standard deviations and standard errors (SE) and used these errors to calculate in quadrature

384   the 2SE uncertainty in the derived TOM concentration. However, as discussed below, the errors in the

385   means do not appear to capture the full variability in Hg(0), particularly at shorter sampling times.

386   3.5.1 September 14th

387   Our most extensive sampling took place on the 14th and we were able to sample for three ~ 2 hour periods

388   between 9 am and 8 pm. On this day there were multiple manifold spikes of $HgBr_2$ and also an Hg(0) spike

389   and we have a made a detailed analysis of the data for each sampling period.

390   The third sampling period which included a large $HgBr_2$ spike provided the only definitive

391   opportunity to demonstrate the capability of 2P-LIF coupled with pyrolysis to measure oxidized mercury.

392   The third sampling period began at ~ hour 17.3 during a manifold $HgBr_2$ spike that began at hour 17. A

393   short Hg(0) spike was also introduced at hour 18. Fig. 7 shows the 2P-LIF signals from the ambient and

394   pyrolyzed sampling lines together with the means and 1 standard deviation. The UM Tekran was offline at

395   this time and so the 2P-LIF concentrations are calibrated by the concentrations reported by the UNR

396   Tekran at the beginning of the Hg(0) spike which are also shown. Both the UNR Tekran and UW Tekran

397   report very similar Hg(0) concentrations during the Hg(0) spike. Both systems report an Hg(0)

398   concentration of 6.7 ng m^{-3} at the beginning of the spike which, since the pre-spike concentration was ~1.9

399   ng m^{-3}, corresponds to a spike concentration of 4.8 ng m^{-3}. This is lower than the calculated spike

400   concentration of 6.1 ng m^{-3} reported by the manifold operators and suggests that the calculated spike was

401   ~27% higher than the actual spike concentration introduced into the manifold. Fig. 8 shows the means of

402   each set of ambient and pyrolyzed measurements together with the 2σ variation and 2SE of the mean. Fig. 9

403   shows the TOM concentrations calculated from the difference together with 2SE in the TOM concentration.

404   The reported spike concentrations and DOHGS measurements are also shown. During the initial sampling

405   period between ~17.3- 17.8 hours the 2P-LIF pyrolysis measurements do not show evidence for an $HgBr_2$

406   spike. Taking the difference between the ambient and pyrolyzed measurements during this period we obtain

407   [TOM] = 0.05±0.05 ng m^{-3}. Shortly before the introduction of the Hg(0) spike we see clear evidence for an

408   increase in the Hg(0) concentration in the pyrolysis sample relative to the ambient sample. We speculate

409   that the manifold adjustments that were made to introduce the additional Hg(0) spike produced either a

410   change in the flow or some other change in the manifold conditions that allowed the $HgBr_2$ spike to reach

411   our pyrolyzer, which, as mentioned above, was located at the manifold. This difference between the two

412   2P-LIF signals is clearly evident by inspection of Fig.7. Fig. 9 shows that the TOM concentration which

413   should consist almost exclusively of $HgBr_2$ is significantly larger than both the reported $HgBr_2$ spike

414   concentration and the concentrations reported by the DOHGS system which are in perfect agreement.

415   Taking the difference between the ambient and pyrolyzed measurements for hour 18.02-18.35 we obtain

416   [TOM] = 1.20±0.17 ng m^{-3} with 2SE uncertainty. It is important to note again that the calculated Hg(0)

417   spike concentration is 27% larger than the measured concentration. This large difference is most likely due

418   to errors in the flows or the permeation source output but it suggests that little confidence can be placed in

419   the calculated concentration of the $HgBr_2$ spike. In addition, it is clear that the DOHGS measurements

420   show a different temporal profile of TOM. The DOHGS system reports TOM concentrations that agree

421   almost exactly with the calculated spike concentration at the beginning of the spike period and drop to a

422   very low background level that is below the detection limit at the end of the reported spike period. In

423   contrast, the 2P-LIF measurements do not show an increased TOM concentration until shortly before the

424   introduction of the Hg(0) spike and they take ~20 minutes to drop to background levels. The UNR

425   speciation systems sample for 1 hour and this is followed by a 1 hour analysis period so they produce a

426   single hourly average every two hours. During this period the UNR speciation system Spec1 sampled for ~

427   20 minutes during the spike period and then for a further 40 minutes. Spec2 was sampling ambient air

428   outside the manifold.

429   SI Figure 4 shows the 7s average of the 2P-LIF signal from the ambient and pyrolysis sample lines

430   for the first sampling period 8-10.45 hours together with the mean and 1standard deviation (1σ) variation

431   in the 2P-LIF signals. SI 5 shows the means together with the 2σ variation and 2SE of the mean. It is clear

432   that there is significant short term variability in the ambient Hg(0) concentration. SI Fig. 6 shows the TOM

433   concentrations calculated from the difference between the pyrolyzed and ambient channels together with

434   the calculated 2SE in the TOM concentration. The reported spike concentration is also shown. If we take

435   the means of the 2P-LIF ambient and pyrolysis measurements during the reported spike period we obtain:

436   ambient: 2.06±0.05 ng m^{-3} and pyrolyzed: 2.21±0.03 ng m^{-3} giving a TOM concentration of 0.145±0.05 ng

437   m^{-3}. The 2P-LIF measurements are consistent with the detection of TOM but they are much lower than the

438   calculated spike and DOHGS measurements shown in Fig. 10.

439   SI Figs.7-9 show the corresponding plots for the second sampling period from ~ 12.2-14 hours.

440   The alternating sampling between the ambient and pyrolysis channels is more even and SI Fig. 7 shows that

441   there is still variability in ambient Hg(0). The means of all the samples give: ambient: 1.72±0.02 ng m^{-3},

442 pyrolyzed: 1.70 ± 0.02 ng m^{-3}. If we take the subset of measurements that coincide with the reported spike

443 we obtain: ambient: 1.79 ± 0.02 ng m^{-3} pyrolyzed 1.77 ± 0.02 ng m^{-3}. In this case, the 2P-LIF measurements

444 do not detect HgBr$_2$ and are not consistent with the reported spike or DOHGS measurements.

445   SI Figs. 10 and 11 show the averages of the TOM concentrations from the 2P-LIF system together

446 with the measurements from the UNR speciation systems, the reported spike concentrations and 5 min

447 DOHGS concentrations. During this sampling period Spec1 sampled from the RAMIX manifold while

448 Spec2 sampled ambient air outside the manifold. Gustin et al.(2013) detail problems with the response of

449 the Spec2 system and applied a 70% correction that is also shown as "Spec2 corrected". Because both the

450 DOHGS and 2P-LIF pyrolysis systems are expected to measure the sum of gaseous (RGM) and particulate

451 (PBM) oxidized mercury we have plotted the sum of the RGM and PBM concentrations from the

452 speciation systems. They are plotted at the mid-point of the 1 hour sampling period.

453   Over most of the measurement period the 2P-LIF pyrolysis and Spec1 measurements are

454 consistent and lower than the DOHGS measurements. The exception is the large spike in TOM seen by the

455 2P-LIF system at hour 18. The spike occurred during the initial portion of Spec1 sampling and, although it

456 measures an increase in RGM relative to Spec2, the magnitude is not consistent with the 2P-LIF pyrolysis

457 observations.

458

459 **3.5.2 September 13th**

460 September 13th was the first day we were able to sample with the pyrolysis system and we sampled over a

461 period of 5 hours. The only manifold spike during this period was an O$_3$ spike at 1pm that lasted one hour

462 so the speciation instruments were attempting to measure ambient RGM. SI Figure 12 shows averages of

463 TOM concentrations as measured by the 2P-LIF pyrolysis system together with the hourly averages as

464 measured by DOHGS and UNR speciation instruments. The x-axis error shows the duration of the 2P-LIF

465 measurements together with 2SE y-axis error bars. Two of the averages of the 2P-LIF measurement give a

466 physically unrealistic negative concentration suggesting that combining the 2SE errors in the means of the

467 ambient and pyrolyzed channels underestimates the uncertainty in the TOM measurement.

468 **3.5.3 September 15th.**

469   On September 15th we sampled from the trailer roof using the same sampling lines and again

470 alternating between the pyrolyzed and unpyrolyzed channels. SI Figure 13 shows the averages of the 2P-

471 LIF signal from the ambient and pyrolysis channels together with the concentrations measured by the

472 Spec2 system that was sampling ambient air outside the manifold. The concentration obtained from the UM

473 denuder samples described below are also shown. The UW DOHGS and Spec1 systems were sampling

474 from the RAMIX manifold with continuous HgBr$_2$ spiking during this period. We see some evidence for

475 measurable RGM in the first hour of the measurements and this is not seen by Spec 2. Later measurements

476 show no evidence for measurable RGM concentrations.

477 **3.6 Limits of 2P-LIF detection of TOM**

478   As we have noted above, the limit of our detection of TOM depends on the short term variability

479   in the ambient Hg(0) concentration because we use a single fluorescence cell and switch between pyrolysis

480   and ambient channels. We have attempted to give an estimate of the uncertainty by taking two standard

481   errors of the means and combining the errors in quadrature to get an estimate of the uncertainty in the TOM

482   concentration. If the mean of the ambient Hg(0) concentration is not fluctuating significantly on the

483   timescale of channel switching this approach should give an accurate estimate of the uncertainty in TOM.

484   In fact our Hg(0) observations show that the fluctuations in the Hg(0) concentration show a significant

485   diurnal variation, with large fluctuations at night, decreasing over the course of morning hours and being

486   smallest in the afternoon. This can be seen in the long term sampling from September 1st and 2nd and in the

487   observations from September 14th. The observation of statistically significant but physically unrealistic

488   negative TOM concentrations on September 13th may be explained by this. Such an artifact could be

489   produced by contamination in the Teflon valve switching system that alternates the flow to the fluorescence

490   cell. This type of contamination should produce a constant bias that is not actually observed. It appears that

491   the short term variability in Hg(0) concentration produces a small bias in some cases that is not averaged

492   out by switching between the ambient and pyrolyzed channels. For example on September 13th the initial

493   sample period of 1.2 hours gives an RGM concentration of 0.06 ± 0.10 ng m^{-3} while two shorter sampling

494   periods at hour 10.5 (36 min sample) and 13.5 (12 min sample) give 0.15 ± 0.09 ng m^{-3}. Our results suggest

495   that the use of single detection channel with switching between ambient and pyrolyzed samples is not

496   adequate to resolve the small concentration differences that are necessary to be able to monitor ambient

497   TOM. It is necessary to set up two detection systems, one continuously monitoring ambient Hg(0) and the

498   other continuously monitoring a pyrolyzed sample stream giving TM, to get the precision necessary to

499   monitor ambient TOM. Over most of the measurement periods our results are consistent with the lower

500   TOM values reported by the UNR speciation instruments although there is a large uncertainty in the

501   concentrations that is actually difficult to quantify. In addition, it is important to emphasize that this was

502   our first attempt to use the pyrolysis approach to attempt to measure TOM. It is possible that the pyrolyzer

503   was not working efficiently on September 13th. The results from September 14th are more difficult to

504   rationalize. The 2P-LIF pyrolysis system has the sensitivity to detect the much higher values of RGM

505   reported by the DOHGS system and the reported spike concentrations of HgBr$_2$. At higher concentrations,

506   as shown in Fig. 9, the 2P-LIF system can monitor HgBr$_2$ with ~10 minute time resolution. Our results,

507   however, cannot be reconciled with those reported by the DOHGS system or the spike concentrations

508   reported by the UW manifold team.

509

510   **3.7 Manual Denuder Measurements:**

511   As we describe above, our use of manual denuders was similar to that described by Landis et al. (2002)

512   with the exception that we did not incorporate the integrated elutriator/acceleration jet and impactor/coupler

513   on the denuder inlet and the denuders were not heated. Landis et al. (2002) suggest that HgCl$_2$ is

514   quantitatively transported through the manual denuder elutriator/impactor inlet when properly heated. In

515   later work Feng et al. (2003) reported that such impactors could reduce the efficiency of RGM collection

516   although in that work there is no reference to the temperature of the impactor. In this work no type of

517   particle filtering was used on the inlets. In addition, we used single photon LIF to monitor the evolution of

518   Hg(0) in real-time as the RGM decomposed on the hot denuder surface during oven analysis. The analysis

519   was carried out in He buffer gas and the Hg(0) concentration was calibrated by manual injections. The first

520   series of measurements, i.e. September 6-14th involved single denuder sampling. On the 15 and 16th we

521   employed tandem sampling with two denuders in series to assess the extent of RGM "bleedthrough". We

522   used two sets of denuders on the 15th and four sets of denuders on the 16th. Fig. 10 shows the raw data for a

523   denuder analysis showing the preheat Hg(0) calibration injections and the temporal profile of the Hg(0) LIF

524   signal for one of the September 16th samples, denuder 1. The two traces correspond to the two monitoring

525   PMTs set at different gains to increase the dynamic range of the measurements. Fig. 11 shows the

526   calibrated profile for the same denuder together with the "blank" i.e. the trailing denuder. The complete set

527   of manual denuder data together with corresponding values for the UNR speciation units that are closest in

528   sampling time are shown in Table 1. Sampling occurred on denuders 1, 4, 6 and 7. The "trailing" denuders

529   which we have treated as blanks, are denuders 3, 5, 8 and 9. The advantage of monitoring the RGM

530   decomposition in real-time is shown in the September 16th data. The temporal decomposition profiles

531   (TDP) for three of the denuders shown in Fig 11 and SI Figures 14 and 15 show reasonable agreement both

532   in absolute concentration of Hg(0) and the time for decomposition to occur. The fourth denuder sample, SI

533   Fig. 16, is a factor of 4-5 higher in concentration and decomposes on a longer time scale with significant

534   structure in the TDP. Comparing the TDPs for all eight denuders it is clear that the TDP for denuder 7,

535   which shows the anomalously high value, is very different from the TPDs for the other three sample

536   denuders. We believe that this TDP is associated with particulate mercury that has impacted on the denuder

537   wall and decomposes on a slower timescale giving a very different temporal profile from RGM that was

538   deposited on the denuder wall. SI Table 1 shows the values of RGM obtained from denuder analysis

539   together with an indication of impact from a PBM component. We have also included measurements from

540   the UNR speciation systems that overlap with, or are close to, the times when our measurements were

541   made. We draw several conclusions from the measurements. The values we obtain from simultaneous

542   measurements that are not influenced by the presence of PBM agree reasonably well with each other, are

543   broadly consistent with the values reported by the Tekran speciation systems and are typically much lower

544   than the values from the UW DOHGS system. Two sets of tandem denuder measurements from September

545   15 and 16 indicate that there is not a significant level of "bleedthrough" onto the trailing denuders. This

546   suggests that the large differences between the DOHGS system and the UNR speciation systems are not

547   due to specific problems with the RAMIX manifold or the speciation systems deployed at RAMIX even

548   though Spec 2 was not functioning properly as documented by Gustin et al. (2013). The tandem sampling

549   also demonstrates that any denuder artifact is not a result of some type of "bleedthough" artifact that is

550   preventing RGM from being quantitatively captured by the first denuder. These results are consistent with

551   prior work by Landis et al. (2002) and Feng et al. (2003). It is also noteworthy that the manually sampled

552   denuders were at ambient temperature in contrast to the speciation denuders that are held at 50 C. Hence

553   the absolute sampling humidities are similar but the relative humidities are very different. Finally, we

554   suggest that there is value in monitoring RGM decomposition in real time as diagnostic of particulate

555   impact when utilizing the annular denuders without the impactor inlet designed to remove coarse

556   particulate matter that may be retained due to gravitational settling

557   **4.0 Implications of RAMIX results.**

558   We think a realistic assessment of the RAMIX results is imperative because the interpretation of

559   the RAMIX data and the conclusions presented by Gustin et al. (2013) and Ambrose et al. (2013) have

560   enormous implications for both our understanding of current experimental approaches to atmospheric

561   sampling of mercury species and to the chemistry itself. Speciation systems using KCl denuder sampling

562   are widely used in mercury monitoring networks worldwide to measure RGM concentrations and the

563   Gustin et al. (2013) and Ambose et al. (2013) papers suggests these results greatly underestimate RGM

564   concentrations with no clear way to assess the degree of bias.

565   **4.1 Intercomparison of Hg(0)**

566   The assessment of the Hg(0) measurements is a little different in the two manuscripts with

567   Ambrose et al. (2013), noting that "comparisons between the DOHGS and participating Hg instruments

568   demonstrate good agreement for GEM" where GEM refers to Hg(0), and they found a mean spike recovery

569   of 86% for the DOHGS measurements of Hg(0), based on comparisons between measured and calculated

570   spike concentrations. Gustin et al. (2013) suggest that the UM Tekran agreed well with measurements of

571   TM reported by the DOHGS system and they "hypothesize that the long exposed Teflon line connected to

572   the UM Tekran unit provided a setting that promoted conversion of RM to GEM, or that RM was

573   transported efficiently through this line and quantified by the Tekran system. The latter seems unlikely

574   given the system configuration…", where RM refers to reactive mercury. As we note above, we believe

575   that the best explanation for discrepancies between the UM and UNR Tekrans is an experimental issue with

576   the UM Tekran response during the initial period of sampling. We would suggest that data from September

577   5th, one of the few occasions when data from multiple instruments agreed over an extended period is not

578   compatible with either transmission or inline reduction of RGM in our sampling line. What is also

579   significant from this data is the very large discrepancy between the spike concentrations as measured

580   independently by three different Tekran systems and confirmed by the relative response of the 2P-LIF

581   measurements and the calculated spike concentration. The discrepancy, on the order of 25-30%, is larger

582   than the manifold uncertainties suggested by Finley et al. (2013). We note other examples of the measured

583   Hg(0) spikes being significantly lower than the calculated concentrations. In prior work we have shown

584   that both the Tekran and 2P-LIF systems show excellent agreement over more than 3 orders of magnitude

585   in concentration when monitoring the variation in Hg(0) in an N_2 diluent. It is to be expected therefore that

586   the "recovery" of high concentration spikes should show good agreement between the different instruments

587   as observed in the September 5th data. The difference between the observations and the calculated manifold

588   spike concentrations is, we would suggest, a reflection of the significant uncertainty in the calculated

589   manifold spike concentration and is not a reflection of reactive chemistry removing Hg(0). In addition,

590   random uncertainties in the flow calculations should not produce a consistently low bias relative to the

591   calculated spike concentrations. As we note above in section 3.1 Ambrose et al. report an increase in the

592   output of their Hg(0) permeation tube after the move to the RAMIX site but this assumes that their Tekran

593   calibration is accurate. The results are consistent with their Tekran measuring too high an output from the

594   permeation device. This is significant if the same Tekran is being used to calibrate the output of the HgBr$_2$.

595   A more difficult issue is the question of resolving the differences in the temporal variation of

596   ambient Hg(0) at the 5 minute timescale as captured by the different instruments. The Tekran systems

597   should be in agreement with a precision of better than 5% and the 2P-LIF system, with a much faster

598   temporal resolution and detection limit, should be capable of matching this. The differences here are not

599   consistently associated with a single instrument with, for example, the 2P-LIF having some systematic

600   offset with respect to the CVAFS systems. The extent to which the larger (i.e. larger than 5%) observed

601   discrepancy which ranged from 10% to 25% is a result of interferences or simply a reflection of instrument

602   precision is difficult to assess. We note again that the UM instruments had to sample through a very long

603   sampling line and we expect that oxidized mercury is deposited on the sampling line. However it is not

604   possible to assess the extent to which oxidized mercury is reduced back to its elemental form introducing

605   small artifacts. As we suggest below, an intercomparison of instrument response to variation in Hg(0)

606   concentrations in a pure N$_2$ diluent with the Hg(0) concentration varying between 1-3 ng m^{-3} would provide

607   a definitive baseline measurement of the instrument intercomparison precision and accuracy. We suggest

608   that such a measurement is a critical component of any future intercomparrison of mercury instrumentation.

609

610   **4.2 Comparison of Total Oxidized Mercury**

611   To the best of our knowledge RAMIX is the only experiment that has measured ambient TOM using

612   multiple independent techniques. It should again be emphasized that the TOM measurements using

613   pyrolysis with 2P-LIF detection were the first attempt to perform such measurements and the use of a

614   single channel detection system introduced large uncertainties into the measurements. The very large

615   discrepancies between the measurements of TOM reported by the DOHGS system, the Tekran speciation

616   systems and the limited number of 2P-LIF pyrolyzer measurements are the most problematic aspect of the

617   RAMIX measurement suite. Work prior to RAMIX and suggested a potential ozone and/or humidity

618   interference in the operation of KCl coated annular denuders and a number of studies since have also

619   reported such an effect (Lyman et al., 2010; McClure et al., 2014). Typically however the differences

620   between the RAMIX measurements are large and are not germane to the differences between the DOHGS

621   and 2P-LIF pyrolyzer measurements. The SI Figures give an example of the differences between the

622   DOHGS measurements and the denuder and 2P-LIF measurements. Ambrose et al. (2013) note that the

623   DOHGS measurements were, on average, 3.5 times larger than those reported by the Spec1 system and

624   summarize the comparison with denuder measurements as follows: "These comparisons demonstrate that

625   the DOHGS instrument usually measured RM concentrations that were much higher than, and weakly

626 correlated with those measured by the Tekran Hg speciation systems, both in ambient air and during HgBr$_2$

627 spiking tests." The discrepancy of a factor of 3.5 is an average value but, for example, examining the

628 September 14 data at ~5 am the DOHGS system is measuring in excess of 500 pg m^{-3} compared with ~20

629 pg m^{-3} measured by the speciation systems, a factor of 25 difference. At this point the Hg(0) concentration

630 was ~ 3 ng m^{-3} so based on the DOHGS measurements oxidized mercury is ~ 15% of the total mercury

631 concentration. A recent study by McClure et al. (2014) provided a quantitative assessment of the extent to

632 which ozone and humidity impact the recovery of HgBr$_2$ on KCl recovery. They note that although they

633 provide a recovery equation to compare with other studies, they do not recommend use of this equation to

634 correct ambient data until more calibration results become available. In Fig 12, we show the ozone

635 concentration and absolute humidity for a 35 hour sampling period on September 13th and 14th that included

636 two ozone spikes and only sampled ambient TOM. Fig 13 shows the expected denuder recovery based on

637 the formula determined by McClure et al. which varies between a typical value of ~70% dropping to ~50%

638 during the ozone spikes. The figure also shows the reported recoveries i.e. the ratio of RGM as measured

639 by either the UNR speciation systems or the 2P-LIF system divided by the value reported by the DOHGS

640 system. These values are typically much lower than those predicted by the McClure recovery expression. In

641 addition, on September 13th and for most of the 14th the 2P-LIF pyrolysis system sees little or no evidence

642 for high spike concentrations of HgBr$_2$ but records levels that fluctuate around those reported by the

643 speciation systems. The one exception is the spike at hour 18 on September 14th.

644   We suggest that the ability of the 2P-LIF pyrolysis system to monitor large spike concentrations is

645 shown by the measurements during the September 14th HgBr$_2$ spike at hour 18. The evidence for an

646 enhancement in the pyrolyzed sample stream is observable in the raw 7s averaged data and becomes clear

647 taking 5 minute averages. The absolute value of the pyrolyzed enhancement is obtained relative to the

648 concentration of the Hg(0) during the spike taken from the measurements by the UNR Tekran that are in

649 excellent agreement with the DOHGS Hg(0) values. The 2P-LIF measurements show a significantly larger

650 HgBr$_2$ concentration and a different temporal profile compared with the DOHGS instrument. In particular,

651 it is very difficult to rationalize the difference between the 2P-LIF and DOHGS systems during the first

652 hour of the spike. We would suggest it is difficult to make the case that both instruments are measuring the

653 same species. It is clear that the 2P-LIF pyrolyzer is operating efficiently based on the clear observation of

654 TOM at the end of the spike. We again note that the 2P-LIF system is not sensitive to TOM. It is important

655 to note that the DOHGS instrument requires an inline RGM scrubber to remove RGM before the

656 measurement of Hg(0). This inline scrubber utilizes deposition on uncoated quartz wool and the results of

657 Ambrose et al. (2013) imply that while uncoated quartz captures RGM efficiently in the presence of O$_3$,

658 quartz with a KCl coating promotes efficient reduction to Hg(0).

659   It is also reasonable to question the extent to which the Tekran speciation systems operated at

660 RAMIX reflect the performance of these systems when normally operated under recommended protocols.

661 As noted above, the operation of the RAMIX manifold and the Tekran speciation systems has been

662 questioned by Prestbo (2014). In our view the two most significant issues are the performance of the two

663 2537 mercury analyzers associated with each speciation system and the reduced sampling rate. The

664 performance of the two 2537 units is detailed in Gustin et al. (2013) and, as they noted, there was a

665 significant response in each instrument. Examination of Fig SI 6 of Gustin et al. (2013) shows the relative

666 responses of the two instruments and, using concentrations up to 25 ng m^{-3} i.e. manifold spikes, they list a

667 regression of 0.72 [Hg(0)] + 0.08 whereas for the non-spike data they obtain 0.62[Hg(0)] + 0.25. Their

668 Table SI 5 lists the regression including spikes as 0.7 (±0.01) + 0.2, with all concentrations expressed in ng

669 m^{-3}. When considering the use of these analyzers to monitor oxidized mercury the important factor to

670 consider is the loading on the gold cartridge. Table SI 3 lists the mean RGM concentrations from manifold

671 sampling as 52 pg m^{-3} for SPEC1 and 56 pg m^{-3} for SPEC2. For a 1 hour sample at 4 L min^{-1} this

672 corresponds to a cartridge loading of 13 pg. This is similar to the cartridge loading for sampling a

673 concentration of 0.6 ng m^{-3} at 4 l min^{-1} for 5 minutes. If we examine Fig SI 6 of Gustin et al. (2013) we see

674 that the regression analyses are based on higher concentrations than 0.6 ng m^{-3}, i.e. higher cartridge

675 loadings. At concentrations of 0.6 ng m^{-3} the ratio of SPEC2:SPEC1 obtained from these regressions would

676 be 1.05, 0.85 and 1.06 depending on which regression formula is used. We should note that based on Table

677 SI 6 the median RGM concentrations in manifold sampling were 41 and 46 pg m^{-3}. The RGM

678 concentrations for free standing sampling were even lower with means of 26 and 19 pg m^{-3} and medians of

679 23 and 14 pg m^{-3} for SPEC1 and SPEC2 respectively. For concentrations below 40 pg m^{-3} the cartridge

680 loading drops below 10 pg and in addition, the Tekran 2537 integration routine becomes significant.

681 Swartzendruber et al. (2009) reported issues with the standard integration routine and note that below

682 cartridge loadings of 10 pg the internal integration routine produces a low bias in the Hg(0) concentration.

683 They recommend downloading the raw data, i.e. PMT output and integrating offline. This issue has

684 recently been discussed by Slemr et al. (2016) in a reanalysis of data from the CARIBIC program. This

685 compounds the problem of correcting the bias between SPEC1 and SPEC2. Because the speciation

686 instruments were sampling at 4 L /min rather than the recommended 10 L/min a large number the

687 measurements made by the speciation systems are based on uncorrected cartridge loadings of less than 10

688 pg m^{-3}. Based on the above we caution against drawing significant conclusions based on differences

689 between SPEC1 and the corrected SPEC2. These differences are the basis of the conclusions of Gustin et

690 al. (2013) that "On the basis of collective assessment of the data, we hypothesize that reactions forming

691 RM (reactive mercury) were occurring in the manifold" (Gustin et al. (2013) abstract). Later they state

692 "The same two denuders, coated by the same operator, were used from Sept 2 to 13, and these were

693 switched between instruments on September 9. Prior to switching the slope for the equation comparing

694 GOM as measured by Spec 1 versus Spec 2 adjusted was 1.7 (r2=0.57, p<0.5, n=76) after switching this

695 was 1.2 (r2=0.62, p<0.05, n=42). This indicates that although there may have been some systematic bias

696 between denuders SPEC 2 adjusted consistently measured more GOM than SPEC 1. We hypothesize that

697 this trend is due to production of RM in the manifold (discussed later)." If reactions in the manifold were

698 producing RM then this production would surely have resulted in the DOHGS measuring artificially high,

699 i.e. higher than ambient, concentrations of oxidized mercury. However, the paper by Ambrose et al. (2013)

700 (written by a subset of the authors of Gustin et al.(2013)) makes no mention of manifold production of

701 oxidized mercury. In fact Ambose et al. (2013) state, in the supplementary information to their paper, "The

702 same two denuders, prepared by the same operator, were used in the Tekran® Hg speciation systems from

703 2 to 13 September. The denuders were switched between Spec. 1 and Spec. 2 on 9 September. From 2 to 9

704 September, the Spec. 1-GOM/Spec. 2-GOM linear regression slope was 1.7 ($r2 = 0.57$; $p < 0.05$; $n = 76$);

705 from 9 to 13 September the Spec. 1-GOM/Spec. 2-GOM slope was 1.2 ($r2 = 0.62$; $p < 0.05$; $n = 42$). These

706 results suggest that the precisions of the GOM measurements made with Spec. 1 and Spec. 2 were limited

707 largely by inconsistent denuder performance."

708   The oxidized mercury concentrations presented by Ambrose et al. (2013) for the RAMIX

709 measurements suggests a well-defined diurnal profile that peaks at night. It is important to note that the

710 error bars on this profile (Figure 3 of Ambose et al.) are one standard error rather than one standard

711 deviation. The standard deviations, which actually give an indication of the range of concentrations

712 measured show much larger errors indicating significant day to day variation in these profiles.

713 Nevertheless, the measurements show much larger oxidized mercury concentrations than the speciation

714 systems and the very limited number of 2P-LIF measurements. As we note below, there is no known or

715 hypothesized chemistry that can reasonably explain the large RGM concentrations seen by the DOHGS

716 instrument. Both Gustin et al. (2013) and Ambrose et al. (2013) draw some conclusions about the chemistry

717 of mercury that have significant implications for atmospheric cycling. Gustin et al. suggest in their abstract

718 that "On the basis of collective assessment of the data, we hypothesize that reactions forming RM were

719 occurring in the manifold." Later in a section on "Implications" they conclude "The lack of recovery of the

720 HgBr$_2$ spike suggests manifold reactions were removing this form before reaching the instruments." The

721 residence time in the RAMIX manifold was on the order of 1s depending on sampling point and there is no

722 known chemistry that can account for oxidation of Hg(0) or reduction of RGM on this timescale. We would

723 suggest that the most reasonable explanation of the discrepancies between the various RAMIX

724 measurements includes both instrumental artifacts and an incomplete characterization of the RAMIX

725 manifold. If fast gas-phase chemistry is producing or removing RGM in the RAMIX manifold the same

726 chemistry must be operative in the atmosphere as a whole and this requires that we completely revise our

727 current understanding of mercury chemistry. The discrepancies between the DOHGS and speciation

728 systems are further indication that artifacts are associated with KCl denuder sampling under ambient

729 conditions but we would suggest that RAMIX does not constitute an independent verification of the

730 DOHGS performance and that the 2P-LIF measurements raise questions about the DOHGS measurements.

731   Ambrose et al. (2013) also suggest that the observations of very high RGM concentrations indicate

732 multiple forms of RGM and that the concentrations can be explained by oxidation of Hg(0), with O$_3$ and

733 NO$_3$ being the likely nighttime oxidants. We have discussed these reactions in detail previously (Hynes et

734 al., 2009) and concluded that they cannot play any role in homogeneous gas phase oxidation of Hg(0).

735 Ambrose et al. (2013) cite recent work on this reaction by Rutter et al. (2012) stating that "On the basis of

736 thermodynamic data for proposed reaction mechanisms, purely gas-phase Hg(0) oxidation by either O$_3$ or

737 NO$_3$ is expected to be negligibly slow under atmospheric conditions; however, in the case of O$_3$-initiated

738 Hg(0) oxidation, the results of laboratory kinetics studies unanimously suggest the existence of a gas-phase

739 mechanism for which the kinetics can be treated as second-order." We would suggest that a careful reading

740 of the cited work by Rutter et al. (2013) demonstrates the opposite conclusion. We provide additional

741 discussion of these issues in the SI and again conclude that O$_3$ and NO$_3$ can play no role in the

742 homogeneous gas phase oxidation of Hg(0).

743

744 **5.0 Future Mercury Intercomparisons:**

745   The discrepancies that are discussed above suggest a need for a careful independent evaluation of

746 mercury measurement techniques. The approaches used during the evaluation of instrumentation for the

747 NASA Global Tropospheric Experiment (GTE) and the Gas-Phase Sulfur Intercomparison Experiment

748 (GASIE) evaluation offer good models for such an evaluation. The Chemical Instrument and Testing

749 Experiments (CITE 1-3) (Beck et al., 1987; Hoell et al., 1990; Hoell et al., 1993) were a major component

750 of GTE establishing the validity of the airborne measurement techniques used in the campaign. The GASIE

751 experiment (Luther and Stetcher, 1997; Stetcher et al., 1997) was a ground based intercomparison of SO$_2$

752 measurement techniques that might be particularly relevant to issues associated with mercury measurement.

753 In particular, GASIE was a rigorously blind intercomparison that was overseen by an independent panel

754 consisting of three atmospheric scientists none of whom were involved in SO$_2$ research. We would suggest

755 that a future mercury intercomparison should be blind with independent oversight. Based on the RAMIX

756 results it should consist of a period of direct ambient sampling and then manifold sampling in both reactive

757 and unreactive configurations. For example an unreactive configuration would consist of Hg(0) and

758 oxidized mercury in an N$_2$ diluent eliminating any possibility of manifold reactions and offering the

759 possibility of obtaining a manifold blank response. Such a configuration would allow the use of both

760 denuder and pyrolysis measurements since it is reasonable to conclude, based on the current body of

761 experimental evidence, that denuder artifacts are associated with ambient sampling with water vapor and

762 ozone as the most likely culprits. A reactive configuration would be similar to the RAMIX manifold

763 configuration with atmospheric sampling into the manifold and periodic addition of Hg(0) and oxidized

764 mercury over their ambient concentrations. The combination of the three sampling configurations should

765 enable instrumental artifacts to be distinguished from reactive chemistry in either the manifold itself or, for

766 example, on the KCl denuder.

767 **6.0 Conclusions**

768 We deployed a 2P-LIF instrument for the measurement of Hg(0) and RGM during the RAMIX campaign.

769 The Hg(0) measurements agreed reasonably well with instruments using gold amalgamation sampling

770 coupled with CVAFS analysis of Hg(0). Measurements agreed to 10-25% on the short term variability in

771 Hg(0) concentrations based on a 5 minute temporal resolution. Our results also suggest that the operation of

772 the RAMIX manifold and spiking systems were not as well characterized as Finley et al. (2013) suggest.

773 We find that the calculated concentration spikes consistently overestimated the amount of Hg(0) introduced

774   into the RAMIX manifold by as much as 30%. This suggests a systematic error in concentration

775   calculations rather than random uncertainties that should not produce a high or low bias.

776   We made measurements of TM, and hence TOM by difference, by using pyrolysis to convert

777   TOM to Hg(0) and switching between pyrolyzed and ambient samples. The short term variation in ambient

778   Hg(0) concentrations is a significant limitation on detection sensitivity and suggests that a two channel

779   detection system, monitoring both the pyrolyzed and ambient channels simultaneously is necessary for

780   ambient TOM measurements. Our TOM measurements were normally consistent, within the large

781   uncertainty, with KCl denuder measurements obtained with two Tekran speciation systems and with our

782   own manual KCl denuder measurements. The ability of the pyrolysis system to measure higher RGM

783   concentrations was demonstrated during one of the manifold $HgBr_2$ spikes but the results did not agree with

784   those reported by the UW DOHGS system. We would suggest that it is not possible to reconcile the

785   different measurement approaches to TOM. While there is other evidence that KCl denuders may

786   experience artifacts in the presence of water vapor and ozone the reported discrepancies cannot explain the

787   very large differences reported by the DOHGS and Tekran speciation systems. Similarly, the differences

788   between the DOHGS and 2P-LIF pyrolysis measurements suggest that one or both of the instruments were

789   not making reliable, quantitative measurements of RGM. We suggest that both instrumental artifacts, an

790   incomplete characterization of the sampling manifold, and limitations in the measurement protocols make

791   significant contributions to the discrepancies between the different instruments and that it would be rash to

792   draw significant implications for the atmospheric cycling of mercury based on the RAMIX results. This is

793   particularly true of the RGM results. If one were to conclude that the discrepancies between the DOHGS

794   and speciation systems sampling ambient oxidized mercury are accurate and reflect a bias that can be

795   extrapolated to global measurements then it means that atmospheric RGM concentrations are much higher

796   than previously thought and that we have little understanding of the atmospheric cycling of mercury. What

797   is not in dispute is the urgent need to resolve the discrepancies between the various measurement

798   techniques. The RAMIX campaign provided a valuable guide for the format of any future mercury

799   intercomparison. It clearly demonstrated the need to deploy high accuracy calibration sources of Hg(0) and

800   oxidized mercury, the need for multiple independent methods to measure elemental and oxidized mercury

801   and to clearly characterize and understand the differences reported by instruments that are currently being

802   deployed for measurements.

803

804   **Acknowledgements**

805   This work was supported by NSF Grant # AGS-1101965, a National Science Foundation Major

806   Instrumental Grant (#MRI-0821174) and by the Electric Power Research Institute. We thank Mae Gustin

807   and her research group and Dr Robert Novak for their hospitality, assistance and use of laboratory facilities

808   during the RAMIX intercomparison. We thank Mae Gustin and Dan Jaffe for the use of their RAMIX data

809   for comparison with our results. We thank Eric Prestbo for helpful comments on the manuscript.

810

811

812 **References:**

813 Ambrose, J. L.; Lyman, S. N.; Huang, J.; Gustin, M. S.; Jaffe, D. A.: Fast time resolution oxidized mercury

814 measurements during the Reno Atmospheric Mercury Intercomparison Experiment (RAMIX), Environ.

815 Sci. Technol., 47 (13), 7284–7294; DOI 10.1021/es303916v, 2013.

816

817 Bauer, D.; Campuzano-Jost, P.; Hynes, A. J., Rapid, ultra-sensitive detection of gas phase elemental

818 mercury under atmospheric conditions using sequential two-photon laser induced fluorescence, J. Environ.

819 Monitor.*,* 4, (3), 339-343, 2002.

820

821 Bauer, D., D'Ottone, L., Campuzano-Jost, P. and Hynes, A. J.: Gas Phase

822 Elemental Mercury: A Comparison of LIF Detection Techniques and

823 Study of the Kinetics of Reaction with the Hydroxyl Radical, J. Photochem.

824 Photobio, 57, 247, 2003

825

826 Bauer, D., Everhart, S., Remeika, J., Tatum Ernest, C., and Hynes, A. J.: Deployment of a Sequential Two-

827 Photon Laser Induced Fluorescence Sensor for the Detection of Gaseous Elemental Mercury at Ambient

828 Levels: Fast, Specific, Ultrasensitive Detection with Parts-Per-Quadrillion Sensitivity, Atmos. Meas.

829 Tech., 7, 4251-4265, www.atmos-meas-tech-discuss.net/7/5651/2014/ doi:10.5194/amtd-7-5651-2014,

830 2014.

831

832 Breckenridge, W. H. and Umemoto, H.: Collisional Quenching of Electronically Excited Metal Atoms, in

833 Advances in Chemical Physics: Dynamics of the Excited State, Volume 50 (ed K. P. Lawley), John Wiley

834 & Sons, Inc., Hoboken, NJ, USA. doi: 10.1002/9780470142745.ch5, 2007.

835

836 Beck, S. M.; Bendura, R. J.; Mcdougal, D. S.; et al.: Operational overview of NASA GTE CITE-1 airborne

837 instrument intercomparisons - carbon-monoxide, nitric-oxide, and hydroxyl instrumentation, J. Geophys.

838 Res., 92, 1977-1985, 1987

839

840 Díez, S.: Human health effects of methylmercury exposure, Rev. Environ. Contam. Toxicol., 198,

841 111−132, DOI: 10.1007/978-0- 387-09646-9, 2009.

842

843 Ebinghaus, R.; Banic, C.; Beauchamp, S.; Jaffe, D.; Kock, H.; Pirrone, N.; Poissant, L.; Sprovieri, F.;

844 Weiss-Penzias, P., Spatial coverage and temporal trends of land-based atmospheric mercury measurements

845 in the Northern and Southern Hemispheres. In Mercury fate and transport in the global atmosphere:

846 Emissions, measurements and models, Mason, R.; Pirrone, N., Eds. Springer: New York, NY, 2009.

847

848   Ernest, C.T.; Donohoue, D.; Bauer, D.; Ter Schure, A.; Hynes, A.J. Programmable Thermal Dissociation of

849   Reactive Gaseous Mercury, a Potential Approach to Chemical Speciation: Results from a Field Study,

850   Atmosphere, *5*, 575-596. doi:10.3390/atmos5030575, 2014

851

852   Feng, X.; Lu, J.Y.; Gregoire, D. C.; Hao, Y.; Banic, C. M.; Schroeder, W. H. : Evaluation and application:

853   of a gaseous mercuric chloride source, ,Anal. Bioanal. Chem., 376, 1137-1140, 2003.

854

855   Finley, B. D.; Jaffe, D. A.; Call, K.; Lyman, S.; Gustin, M.; Peterson, C.; Miller, M.; Lyman, T. :

856   Development, testing, and deployment of an air sampling manifold for spiking elemental and oxidized

857   mercury during the Reno Atmospheric Mercury Intercomparison Experiment (RAMIX). Environ. Sci.

858   Technol., 44, 7277-7284,

859   DOI 10.1021/es304185a, 2013

860

861   Gustin, M.; Jaffe, D., Reducing the Uncertainty in Measurement and Understanding of Mercury in the

862   Atmosphere. Environ. Sci. Technol., 44, (7), 2222-2227, 2010.

863

864   Gustin, M. S.; Huang, J.; Miller, M. B.; Peterson, C.; Jaffe, D. A.; Ambrose, J.; Finley, B. D.; Lyman, S.

865   N.; Call, K.; Talbot, R.; Feddersen, D.; Mao, H.; Lindberg, S. E. Do we understand what the mercury

866   speciation instruments are actually measuring? Results of RAMIX. Environ. Sci. Technol., 47 (13), 7295–

867   7306; DOI 10.1021/es3039104, 2013.

868

869   Hoell, J. M. ; Albritton, D. L.; Gregory, G. L.; et al., Operational overview of NASA GTE/CITE-2 airborne

870   instrument intercomparisons - nitrogen-dioxide, nitric-acid, and peroxyacetyl nitrate, J. Geophys. Res., 95

871   10047-10054, 1990

872

873   Hoell, J. M. ; Davis, D. D.; Gregory, G. L.; et al., operational overview of the NASA GTE CITE-3 airborne

874   instrument intercomparisons for sulfur-dioxide, hydrogen-sulfide, carbonyl sulfide,dimethyl sulfide, and

875   carbon-disulfide, J. Geophys. Res., 98, 23291-23304, 1993.

876

877   Houyoux, M.; Strum, M. Memorandum: Emissions Overview: Hazardous Air Pollutants in Support of the

878   Final Mercury and Air Toxics Standard; EPA-454/R-11-014; Emission Inventory and Analysis Group Air

879   Quality Assessment Division: Research Triangle Park, NC, USA, 2011.

880   Available at: http://www.epa.gov/mats/pdfs/20111216EmissionsOverviewMemo.pdf

881

882   Hynes, A. J.; Donohoue, D. L.; Goodsite, M. E.; Hedgecock, I. M., Our current understanding of major

883   chemical and physical processes affecting mercury dynamics in the atmosphere and at the air-

884 water/terrestrial interfaces In *Mercury Fate and Transport in the Global Atmosphere*, Mason, R.; Pirrone,

885 N., Eds. Spring New York, NY, 2009.

886

887 Landis, M. S.; Stevens, R. K.; Schaedlich, F.; Prestbo, E. M., Development and Characterization of an

888 Annular Denuder Methodology for the Measurement of Divalent Inorganic Reactive Gaseous Mercury in

889 Ambient Air. Environ. Sci. Technol., 36, (13), 3000-3009, 2002.

890

891 Lin, C.J.; Pongprueksa, P.; Lindberg, S.E.; Pehkonen, S.O.; Byun, D.; Jang, C. Scientific uncertainties in

892 atmospheric mercury models I: Model science evaluation. Atmos. Environ., 40, 2911–2928,

893 doi:10.1016/j.atmosenv.2006.01.009, 2006.

894

895 Lindberg, S. E.; Bullock, R.; Ebinghaus, R.; Engstrom, D.; Feng, X.; Fitzgerald, W.; Pirrone, N.; Prestbo,

896 E.; Seigneur, C. A synthesis of progress and uncertainties in attributing the sources of mercury in

897 deposition. Ambio, 36, 1932, 2007.

898

899 Luther III, D. W.; D.L.; Stecher, H.A. III, Preface: Historical background. J. Geophys. Res., 102, 16215-

900 16217, 1997.

901

902 Lyman, S. N.; Jaffe, D. A.; Gustin, M. S. Release of mercury halides from KCl denuders in the presence of

903 ozone. Atmos. Chem. Phys. 2010, 10, 8197–8204; DOI 10.5194/acp-10- 8197- 2010.

904 Mason, R.A. Mercury Emissions from Natural Processes and their Importance in the Global Mercury

905 Cycle. In Mercury Fate and Transport in the Global Atmosphere: Emissions, Measurements and Models;

906 Pirrone, N., Mason, R.A., Eds.; Springer: Dordrecht, The Netherlands, 2009; pp. 173–191.

907

908 Mergler, D.; Anderson, H.A.; Chan, L.H.M.; Mahaffey, K.R.; Murray, M.; Sakamoto, M.; Stern, A.H.

909 Methylmercury exposure and health effects in humans: A worldwide concern. Ambio 2007, 36, 3–11,

910 doi:10.1579/0044-7447(2007)

911

912 McClure, C. D.; Jaffe, D.A.; Edgerton, E.S. Evaluation of the KCl denuder method for gaseous oxidized

913 mercury using HgBr2 at an in-service AMNet site. | Environ. Sci. Technol. 2014, 48, 11437−11444,

914 dx.doi.org/10.1021/es502545k

915

916 Prestbo, E. C., Air Mercury Speciation Accuracy and Calibration, SFO Air Mercury Speciation Workshop,

917 July 2014. Available at http://omicsgroup.com/conferences/ACS/conference/pdfs/14958-Speaker-Pdf-T.pdf

918

919 Prestbo, E. C., Chief Scientist, Tekran Instruments, personal communication, 2015.

920

921   Pirrone, N.; Cinnirella, S.; Feng, X.; Finkelman, R.B.; Friedli, H.R.; Learner, J.; Mason, R.; Mukherjee,

922   A.B.; Stracher, G.; Streets, D.G.; Telmer, K. Global mercury emissions to the atmosphere from natural and

923   anthropogenic sources. In Mercury Fate and Transport in the Global Atmosphere: Emissions,

924   Measurements and Models; Pirrone, N., Mason, R.A., Eds.; Springer: Dordrecht, The Netherlands, 2009;

925   pp. 3–50.

926

927   Rutter, A.P., Shakya, K.M., Lehr, r., Schauer, J. J., Griffin, R. J., Oxidation of gaseous elemental mercury

928   in the presence of secondary organic aerosols, Atmos. Environ., 59, 86-92, 2012

929

930   Scheuhammer, A.M.; Meyer, M.W.; Sandheinrich, M.B.; Murray, M.W., Effects of environmental

931   methylmercury on the health of wild bird, mammals, and fish. Ambio 2007, 36, 12–18, doi:10.1579/0044-

932   7447(2007)36[12:EOEMOT]2.0.CO;2.

933

934   Selin, N.E., Global Biogeochemical Cycling of Mercury: A Review. Annu. Rev. Environ. Resour. 2009,

935   34, 43–63, doi:10.1146/annurev.environ.051308.084314.

936

937   Slemr, F., Brunke, E.-G., Ebinghaus, R., and Kuss, J.: Worldwide trend of atmospheric mercury since

938   1995, Atmos. Chem. Phys., 11, 4779-4787, doi:10.5194/acp-11-4779-2011, 2011.

939

940   Slemr, F.; Weigelt, A.; Ebinghaus, R.; Kock, H.H.; Bödewadt, J.; Brenninkmeijer, C.A.M.; Rauthe-Schöch,

941   A. ; Weber, S.; Hermann, M.; Zahn, A.; Martinsson, B.; Atmospheric mercury measurements onboard the

942   CARIBIC passenger aircraft, Atmos. Chem. Phys., accepted for publication, 2016

943

944   Sprovieri, F., Pirrone, N., Bencardino, M., D'Amore, F., Carbone, F., Cinnirella, S., Mannarino, V., Landis,

945   M., Ebinghaus, R., Weigelt, A., Brunke, E.-G., Labuschagne, C., Martin, L., Munthe, J., Wängberg, I.,

946   Artaxo, P., Morais, F., Barbosa, H. D. M. J., Brito, J., Cairns, W., Barbante, C., Diéguez, M. D. C., Garcia,

947   P. E., Dommergue, A., Angot, H., Magand, O., Skov, H., Horvat, M., Kotnik, J., Read, K. A., Neves, L.

948   M., Gawlik, B. M., Sena, F., Mashyanov, N., Obolkin, V., Wip, D., Feng, X. B., Zhang, H., Fu, X.,

949   Ramachandran, R., Cossa, D., Knoery, J., Marusczak, N., Nerentorp, M., and Norstrom, C.: Atmospheric

950   mercury concentrations observed at ground-based monitoring sites globally distributed in the framework of

951   the GMOS network, Atmos. Chem. Phys., 16, 11915-11935, doi:10.5194/acp-16-11915-2016, 2016.

952

953   Streets, D.G.; Devane, M.K.; Lu, Z.; Bond, T.C.; Sunderland, E.M.; Jacob, D.J. All-time releases of

954   mercury to the atmosphere from human activities. Environ. Sci. Technol. 2011, 45, 10485–10491.

955

956   Stecher, H.A. III; Luther III, G.W.; MacTaggart, D.L.; Farwell, S.O.; Crossley, D.R.; Dorko, W.D.;

957   Goldan, P.D.; Beltz, N.; Krischke, U.; Luke, W.T.; Thornton, D.C.; Talbot, R.W.; Lefer, B.L.; Scheuer,

958 E.M.; Benner, R.L.; Wu, J.; Saltzman, E.S.; Gallagher, M.S.: Ferek, R.J.: Results of the gas- phase sulfur

959 intercomparison experiment (GASIE): Overview of experimental setup, results, and general conclusions, J.

960 Geophys. Res., 102, 16219-16236, 1997

961

962 Subir, M.; Ariya, P. A.; Dastoor, A. P., A review of uncertainties in atmospheric modeling of mercury

963 chemistry I. Uncertainties in existing kinetic parameters : Fundamental limitations and the importance of

964 heterogeneous chemistry. Atmos. Environ., 45, 5664-5676, 2011.

965

966 Subir, M.; Ariya, P.A.; Dastoor, A.P. A review of the sources of uncertainties in atmospheric mercury

967 modeling II. Mercury surface and heterogeneous chemistry—A missing link. Atmos. Environ. 2012, 46, 1–

968 10, doi:10.1016/j.atmosenv.2011.07.047.

969

970 Swartzendruber, P. C.; Jaffe, D. A.; Finley, B.:Improved fluorescence peak integration in the Tekran 2537

971 for applications with sub-optimal sample loadings, Atmos. Environ., Volume 43, 3648-3651, 2009

972

973 UNEP, United Nations Environment Program, Chemicals Branch, Mercury, Time to Act, 2013

974

975 UNEP, United Nations Environment Program, Chemicals Branch, The Global Atmospheric Mercury

976 Assessment: Sources, Emissions and Transport, 2008.

977

978 UNEP, United Nations Environment Program: http://www.mercuryconvention.org, 2014

979

980 U.S. EPA, United States Environmental Protection Agency "Mercury Research Strategy", EPA/600/R-

981 00/073, , September 2000.

982

983 U.S. EPA, United States Environmental Protection Agency, Mercury and Air Toxics Standard. Available

984 online: http://www.epa.gov/mats (accessed on 13 March 2013).

985

986

987

988

989

990   Table 1: RAMIX Manual KCl Denuder Sampling

Date	Sample time	mid point	sample	blank		time	spec1		spec2 (uncorr	
	hours	hour	pg m^{-3}	pg m^{-3}			GOM	PBM	GOM	PBM

							pg m-3	pg m-3	pg m^{-3}	pg m^{-3}
9/6	1.5	15	127.9*	2.27		13:00	200.7	51.8	205.1	4.3
						15:00	65.7	32.0	84.9	6.0
9/7	2	16	112.9*	0		14:00	39.8	136.4	94.3	2.5
			21.2			16:00	48.5	177.3	68.9	1.5
			285.8*			18:00	28.1	182.2	37.4	3.3
			30.6							
9/10	3	15.3	74.3	1995		14:00	26.7	10.5	27.4	4.2
			44.2			16:00	24.1	18.3	23.7	2.3
9/13	4	15	12.8	8.2		13:00	0.7	16.9	0.5	16.6
			13.56			17:00	37.6	16.1	25.2	2.7
9/14	4.5	14	39*	3.3		12:00	34.9	12.0	23.9	5.5
			17.3			14:00	57.	18.4	26.3	38.6
						16:00	42.0	17.4	26.3	4.0
9/15	4.5	15	15.24	1.53		13:00	113.9	39.1	27.6	3.9
			20.4	4.87		15:00	80.6	22.2	17.7	3.9
						17:00	110.8	24.1	8.6	8.1
9/16	2.75	16	148*	5		8:00	19.7	4.7	14.8	5.4
			42	6		9:00				
			26	5		10:00	28.7	13.3	19.9	4.8
			47	4						

991   •   * evidence from TDP's for presence of PBM

992   •   Measurements for UNR Speciation system made at similar times. The Spec 2 measurements are

993   uncorrected values.

994

995

996

997

998

999

1000

1001

1002

1003

1004

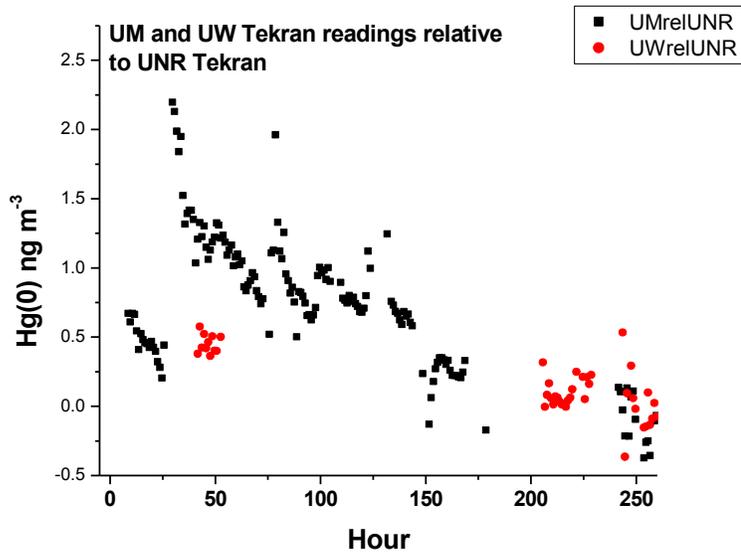
1005

1006

1007

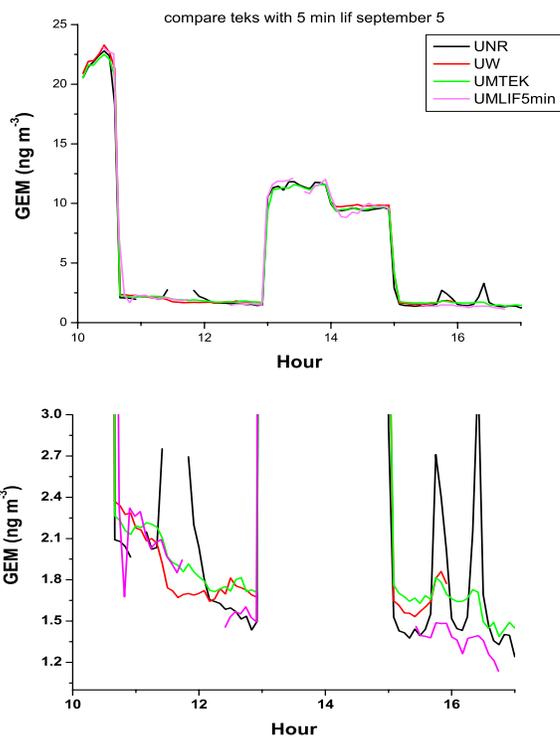
1008

1009

1010 **Figures.**

1011

1012 Figure 1. Comparison of Hg(0) readings from the UM, UW and UNR Tekrans over the first 260 hours of

1013 UM measurements. The absolute concentration difference relative to the UNR instrument is shown in black

1014 for the UM Tekran and in red for the DOHGS (UW) Tekran.

1015

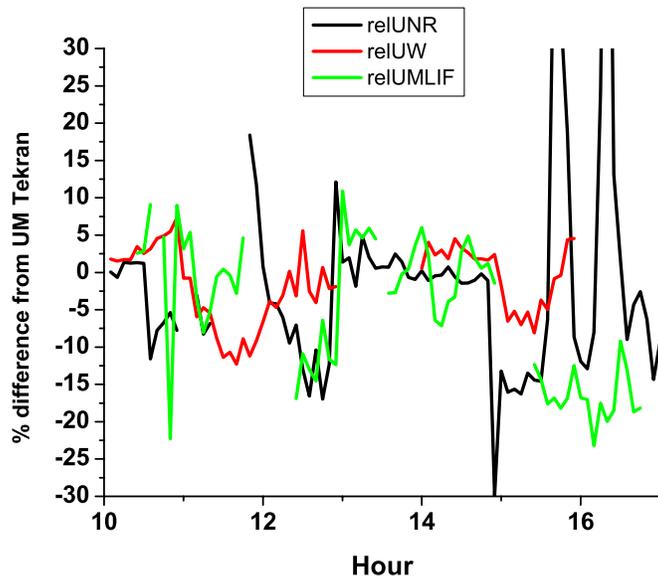
1016   Figure 2: a). A seven hour sequence of GEM measurements from September 5th that included two manifold

1017   spikes. Shown are the sequence of GEM measurements from the UNR, UW and UM Tekrans together with

1018   the 5 minute averages of the 2P-LIF signal. b) An expanded concentration scale focusing on ambient

1019   measurements.

1020

1021

1022

1023

1024   Figure 3: Seven hour measurement period from September 5th. The % difference of the UNR (black line)

1025   and UW (red line) Tekrans and the UM 2P-LIF (blue line) measurements relative to the UM Tekran is

1026   shown.

1027

1028

1029

1030

1031

1032

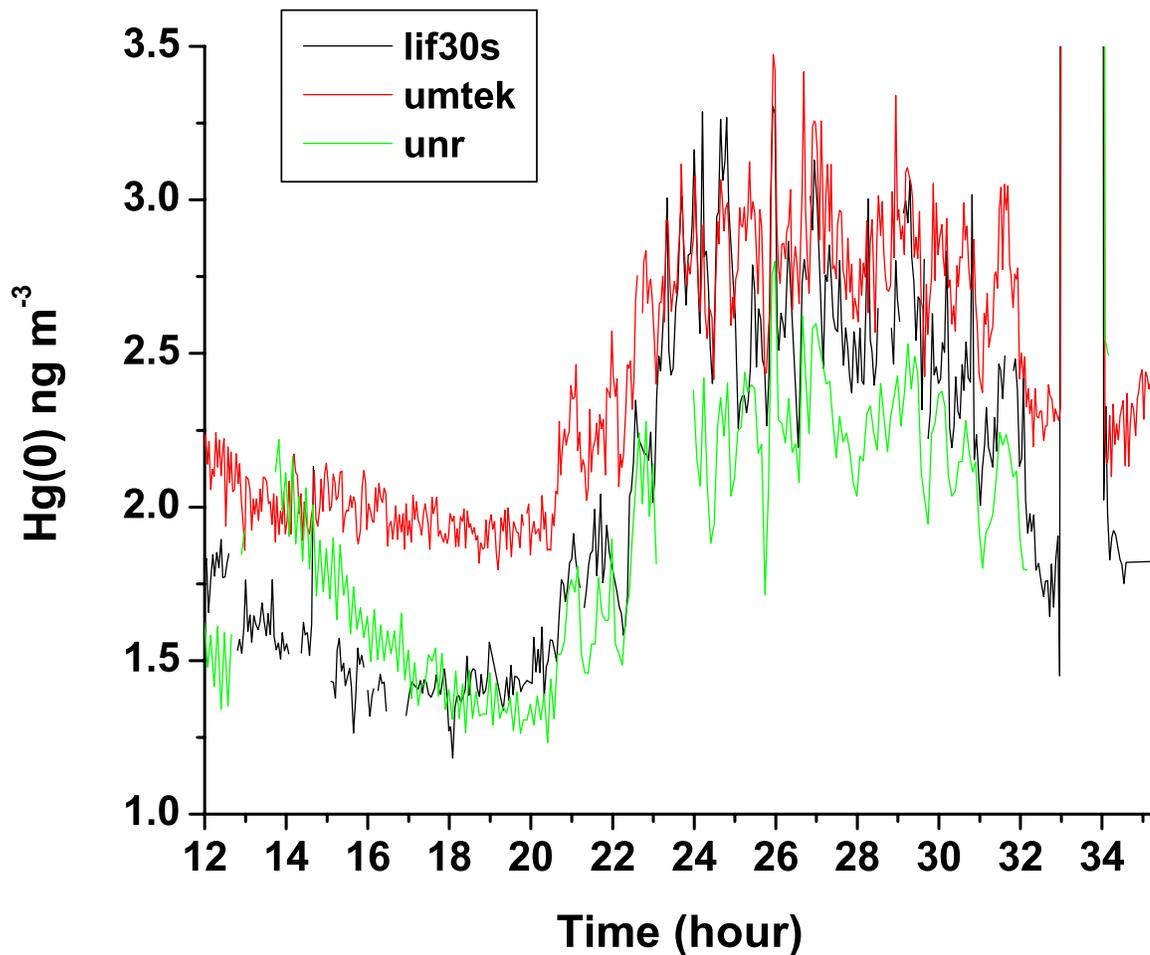
1033

1034

1035

1036

1037

1038

1039
1040
1041
1042 Fig 4: 22 hour sampling period from September 1st and 2nd. Comparison of the UM (red line) and UNR
1043 (green line) Tekrans with the UM 2P-LIF (black line) concentrations. The concentrations for each
1044 instrument are scaled to force agreement during the second manifold spike at hour 33. This is the data from
1045 SI Fig. 3 with the concentration scale expanded to shown only ambient data.
1046
1047
1048
1049
1050
1051
1052

1053
1054
1055
1056
1057
1058   Fig 5: A section of the 22 hour sampling period from September 1st and 2nd. Comparison of the UM (red

1059   line) and UNR (green line) Tekrans with the UM 2P-LIF (black line) concentrations. The concentrations

1060   for each instrument are scaled to force agreement during the second manifold spike at hour 33. This is the

1061   data from SI Fig. 3 with the concentration scale expanded to shown only ambient data between hours 29

1062   and 32.

1063
1064

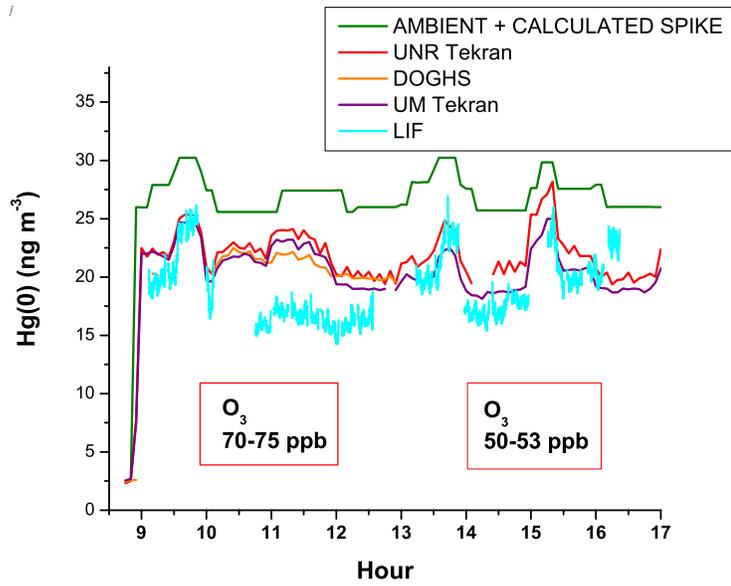
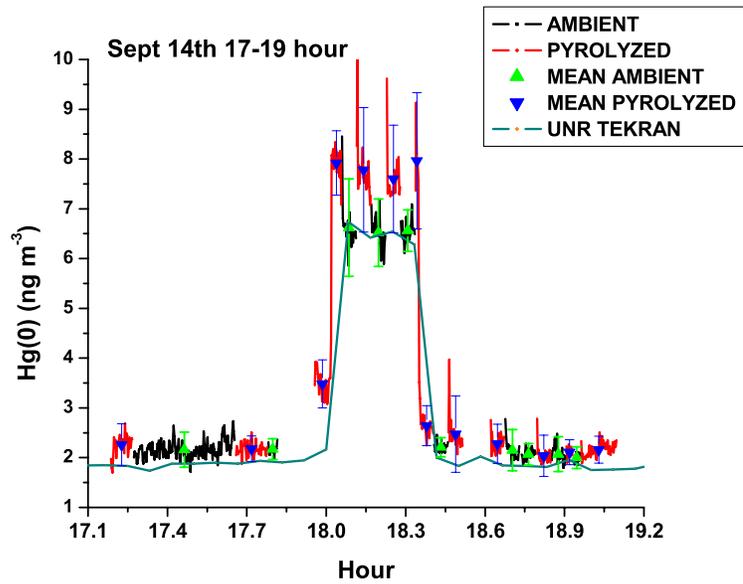



Figure 6: September 7th an ozone interference test. A comparison of the UM, UW and UNR Tekrans and the UM-2P-LIF measurements. The "expected" concentration calculated from the ambient Hg(0) concentration prior to the spike plus the calculated spike concentration is also shown.

1072

1073

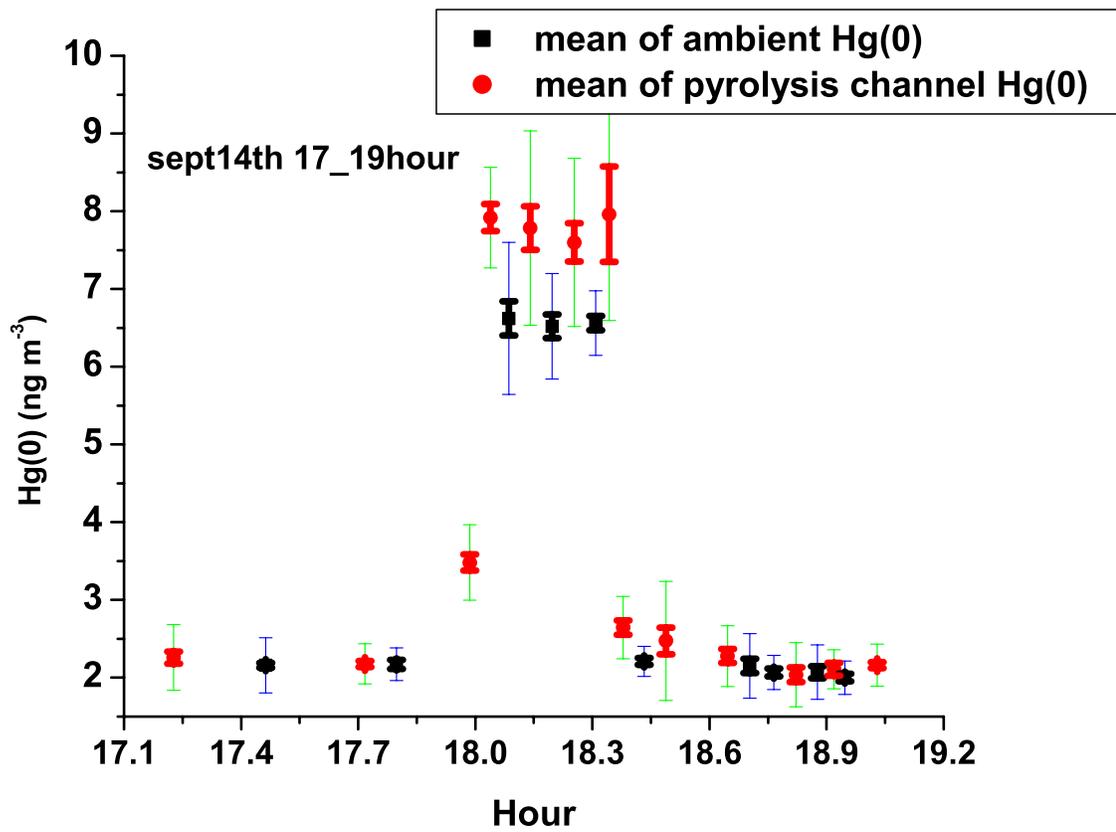
1074   Fig 7: September 14 measurements hours 17-19 (5-7pm). The background subtracted 2P-LIF signals from

1075   the ambient (black) and pyrolyzed sampling lines (red) are shown. The gaps correspond to times when the

1076   laser was blocked to check power and background. The means and 1 standard deviation of each sample are

1077   shown. The absolute Hg(0) concentrations are obtained by scaling the ambient Hg(0) signal to the absolute

1078   Hg(0) concentration reported by the UNR Tekran during the Hg(0) manifold spike.

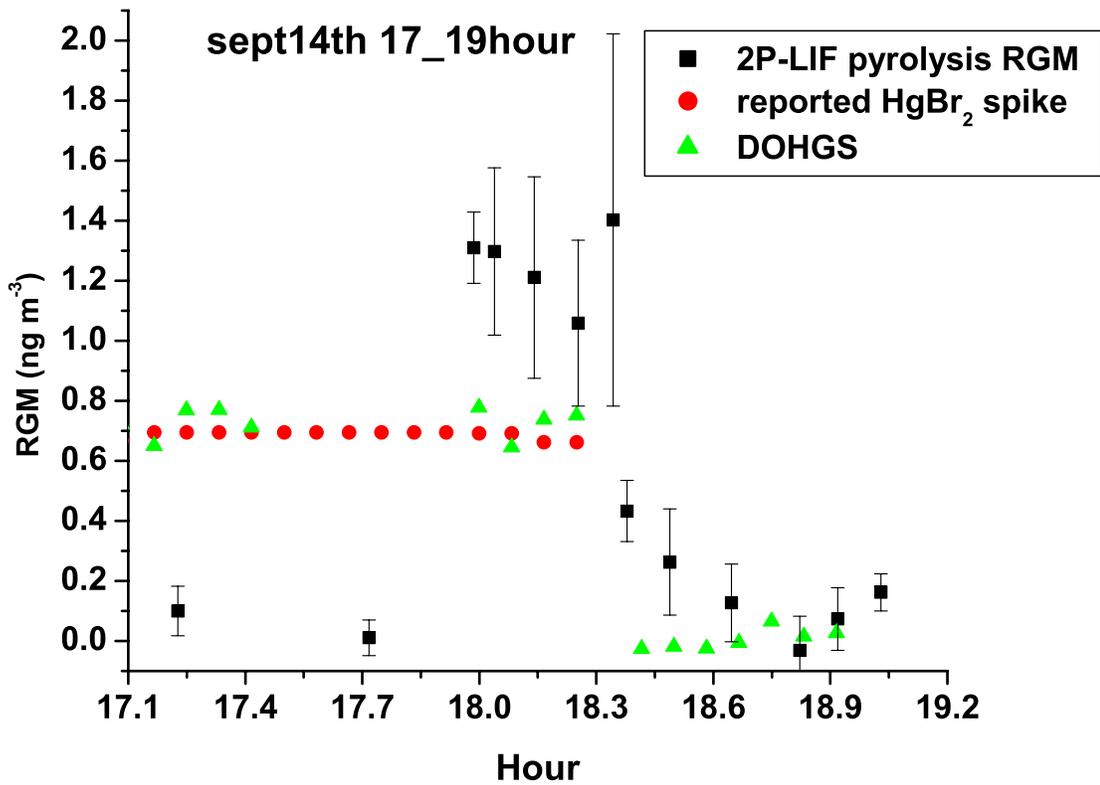
1079

1080

1081

Fig. 8: September 14 measurements hour 17-19. The means of the ambient channel (black) and pyrolyzed

channel (red) are shown. The error bars show both 2 standard errors (thicker line) and 2 standard

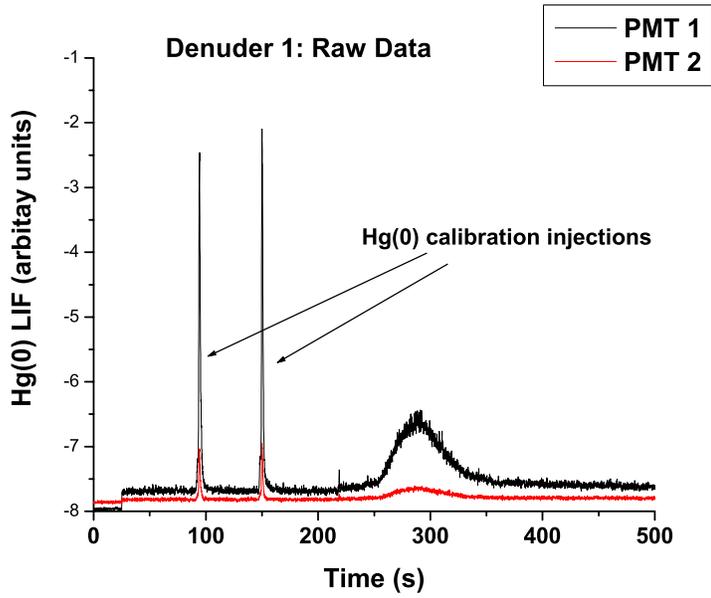
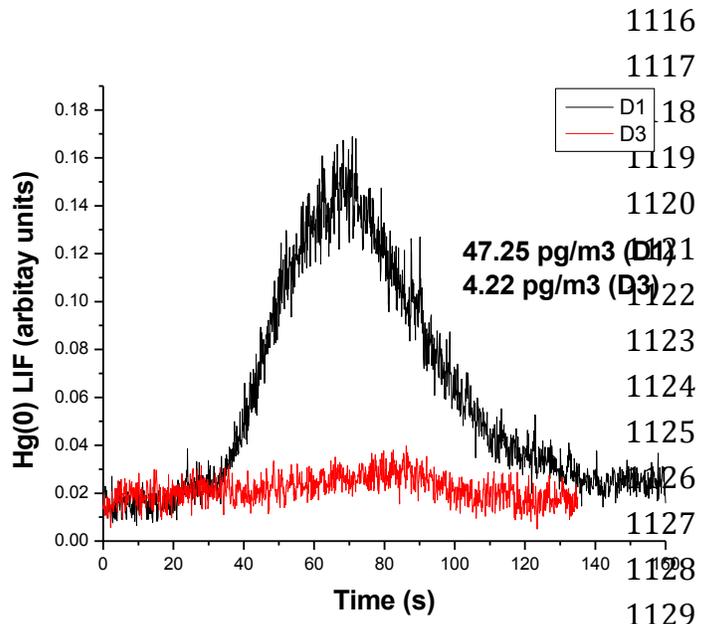
deviations.

1085

1086
1087
1088 Fig 9: TOM concentrations calculated from the difference between the pyrolyzed and ambient sample

1089 concentrations together with 2SE in the TOM concentrations. The reported $HgBr_2$ spike concentrations and

1090 DOHGS measurements are also shown.

1091
1092

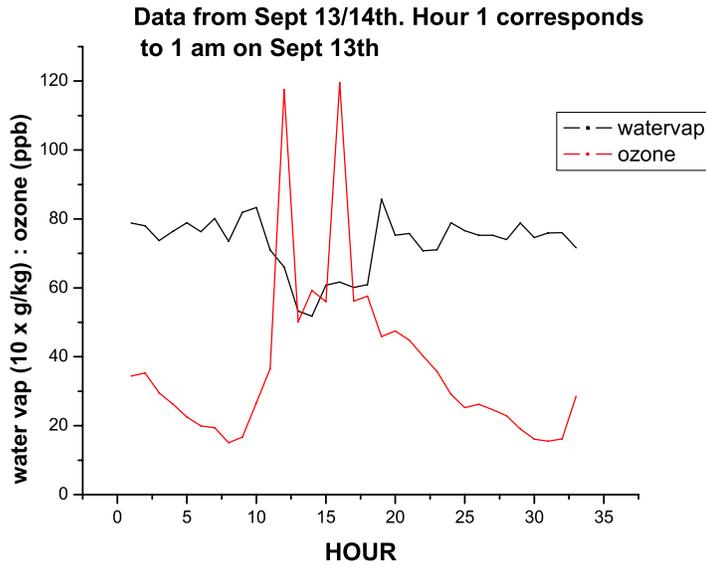
1093

1094
1095
1096   Fig. 10: September 16th KCl manual denuder measurements. The raw data for the temporal decomposition
1097   profiles (TDP) for the denuder D1 is shown.
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115

1130   Fig. 11 September

1131   16th KCl manual

1132   denuder measurements. The calibrated temporal decomposition profiles (TDP) for the tandem denuder pair,

1133   D1 and D3 are shown.

1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152

1153

1154
1155
1156
1157   Fig. 12: The ozone concentration and absolute humidity for a 35 hour sampling period on September 13th

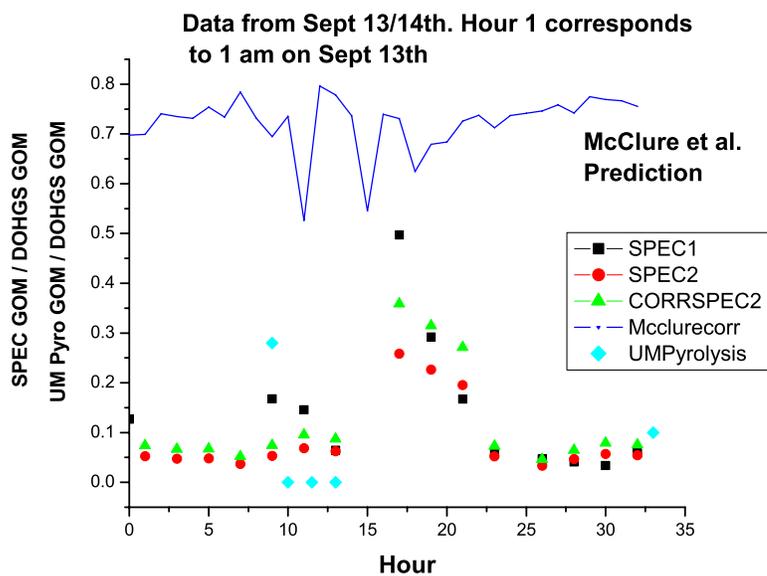
1158   and 14th that included two ozone spikes and only sampled ambient TOM.

1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173

1174
1175
1176   Fig. 13. Expected denuder recovery based on the formula determined by McClure et al. which varies

1177   between a typical value of ~70% dropping to ~50% during the ozone spikes. The figure also shows the

1178   reported recoveries i.e. the ratio of RGM as measured by either the UNR speciation systems or the 2P-LIF

1179   system devided by the value reported by the DOHGS system.

1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197

1198
1199
1200