

Response to Reviewers

Weiss-Penzias et al. “Use of a global model to understand speciated atmospheric mercury observations at five high-elevation sites”

Reviewer #1

Comment #1: The paper is difficult to read, especially for people who are interested in the cycle of atmospheric mercury but are not familiar with the detailed discussion of oxidation mechanisms. A brief presentation of the two oxidation mechanisms and their deficiencies at the outset would help. From such presentation the authors could derive criteria or relationships which could distinguish between the schemes. The search for such criteria or relationships would then provide a backbone for the paper and make it more readable. I still think that the investigations in the paper are valuable and should be published in a final version when the authors improve the readability and take into consideration the following comments and suggestions:

Response to comment 1: We have tried to clarify the take away message of the manuscript and made the arguments easier to follow. We have included the mercury redox reactions that are used by the GEOS-Chem model. We realize that the manuscript raised the idea that the reaction mechanisms were somehow deficient and that by comparing with the observations we would somehow shed light on this. This was not our intention and we are simply comparing the data with the model, run with two different set of assumptions, to see if any patterns would emerge that would point toward areas of future research both on improving the measurements and the models. We have revised the manuscript to make this objective clearer. The other main objective, or “backbone” of the paper is to compare observations across surface sites that receive consistent input from the free troposphere. Since reactive mercury (GOM, PBM) largely comes from the FT at these relatively remote sites comparisons across space and time were made to shed light on atmospheric processes affecting reactive mercury concentrations. We have again, revised the manuscript to make this objective clearer.

Comment #2: The accuracy of RM measurements is adequately discussed in Section 2.2. But how comparable are the GEM measurements? The RAMIX intercomparison showed that the GEM measurements might not be always comparable. The accuracy of GEM measurements is also needed for the comparison of model data with observations.

Response to comment 2: While the RAMIX results for GEM did identify one out of 4 instruments that had a significant discrepancy, 3 out of the 4 instruments had very similar response for GEM (within 10%) (Gustin et al 2013). We have adjusted the manuscript to say that discrepancies in GEM between instruments is normally < 10%, and although it is a limitation that the Tekrans in this paper were not compared side-by-side, all the instruments were run by skilled operators and the data represent the best available for such a comparison study.

Comment #3: The given observed and modelled concentrations probably refer to m3 at 273.14 K and 1013 hPa? This has to be stated in Sections 2.2 and 2.3.

Response to comment 3: The reviewer is correct and this information has been added to the revised manuscript.

Comment #4: The WV criterion works probably well for the mountain top stations but I wonder about its efficiency at DRI and NV02 which are both in desert valleys, with DRI being in the vicinity of substantial anthropogenic emissions of all sorts of pollutants. One would thus expect larger differences between the WV filtered and unfiltered GEM and RM concentrations than those shown in Fig. 1. Could the authors demonstrate the WV criterion efficiency using other pollutants such as CO, NO_y, or concentrations of condensation nuclei? The somewhat higher average GEM concentration in the unfiltered than in WV filtered data at DRI and NV02 alone does not prove that the WV filtered data are free from local influence.

Response to comment 4: We feel that the WV criterion of including data that had WV < 75% percentile keeps a relatively large portion of the observations for statistical power, but removes the most obvious time periods when the sites were receiving air from the boundary layer. Strong relationships between WV and mercury species were demonstrated at the NV02 and DRI sites in previous work (Weiss-Penzias et al., 2009) from summertime data. These sites are both situated in areas with geogenic sources which cause enhancements in GEM concentrations when air has been in contact with the surface as opposed to air that is from the FT. In the present work, we were able to evaluate the WV criterion by comparing NO_x concentrations at NV02 during the summer of 2007 in the WV<75% and WV>25% categories of data. We report in the revised manuscript that NO_x concentrations in drier air averaged 0.12 ppb, and NO_x in moister air averaged 0.53 ppb. There was also a positive correlation between NO_x and GEM with an r² value of 0.57. This suggests the WV criterion was adequate for this site. No criteria pollutant data were available for DRI so this comparison could not be made. However, DRI sits in the middle of a Hg mineralized old mining district and GEM concentrations measured there are higher than that measured at other sites (typically). The site looks over Reno and the prevailing wind is from the W. During the years the data was collected the DRI facility was smaller –only 2 buildings and there was no industry nearby and the city impacts with respect to Hg are minimal (Reno is not an industrial city-but a gaming town). This location (DRI) is primarily in the free troposphere at night due to the collapse of the boundary layer in the valley that can be quite shallow 200m and would be in the boundary layer later in the day when also the water vapor was higher. We agree that this does not prove the WV filtered data are free from local influence but the low water vapor indicates free troposphere input for these (Lyman and Gustin (2008)). The fact the GEM values are smaller at DRI with the WV filtered data show that the WV criteria works because the impact of local natural enrichment is reduced.

Comment #5: Lyman and Jaffe (2011) reported RM/GEM slope in FT air being close to -1 opposite to some -0.5 in the lower stratosphere. Here both the modelled and even more the observed slopes are in absolute terms much smaller and the observed ones are sometimes even positive. These deviations from the intuitive slope of -1, both of the model and the observations deserve a more extended discussion, not only in Section 3.3.

Response to comment 5: Slopes of GOM vs. GEM of around -1 have been reported previously (Swartzendruber et al., 2006; Lyman and Jaffe, 2012). These have been for specific events, when one particular air mass has been measured, and when THg is likely constant. For these conditions, a slope of -1 indicates that photochemical conversion of GEM to GOM has likely occurred and there have been limited losses of GOM due to scavenging and deposition, and limited replenishment of GEM from the background pool. The RM/GEM slopes reported in Figure 2 in the manuscript and Table 2 in the supplemental information are much greater than -1 (or -1000 pg ng⁻¹), in other words, the slopes are less steep and the relationship between RM and GEM is weaker than the ideal -1 slope. The slope at DRI in

the spring was positive. For these data we do not expect a slope of -1 since these are across an entire season. Over such a long time period, GEM concentrations do not stay constant, especially at DRI which has regular inputs from local natural enrichment, scavenging occurs at varying rates, and thus the lifetime of RM is highly variable. Also the collection efficiency of the Tekran may not be constant under all conditions. However, the nearness to -1 in the seasonally derived slopes is an indicator of the frequency with which photochemical conversion in the FT dominates speciated mercury concentrations at the surface and thus we compare the RM/GEM slopes observed and modeled in this paper. The revised manuscript now includes the above discussion for improved clarity.

Comment #6: Page 2274, line 6: “..RM reached 260, 250, and 100 pg m⁻³ ..” is probably correct.

Response to comment 6: The reviewer is correct and this has been fixed in the revised manuscript.

Comment #7: Section 3.3, page 2274, line 20: Are 23 ppb of O₃ the seasonal mean or the unusually low concentration? Please clarify. The O₃ minimum of 23 ppb coinciding with a GEM minimum and RM maximum at MBO shown in Fig. 4 is a clear sign of transport from marine boundary layer which is not captured by the models. Of course such event has nothing to do with FT and has to be excluded from the FT chemistry discussion. What would be the RM/GEM slope without this event? Another problem with the interpretation of MBO observations is that the observed O₃ is substantially lower than the modelled one, opposite to DRI and NV02 sites where they roughly agree. In summary, the discussion in the section 3.3 is not quite correct and has to be fixed.

Response to comment 7: This time period was complex. High RM was first observed in the low O₃ air mass, but then O₃ recovered to more typical values, while RM remained high. Water vapor was relatively low throughout this period. The 23 ppb of O₃ in the early period was unusually low. The revised manuscript has been amended. We interpret the RM event as follows: June 21 brought an air mass to MBO that was transported at low latitudes and was photochemically processed, with a maximum CO concentration of only 63 ppb, maximum σ_{sp} of 1 Mm⁻¹, and the aforementioned O₃ concentration, and labeled as an “MBL” event by Timonen et al. 2013. This event was followed by another RM event on June 22, when O₃ rebounded to 50 ppb which is more characteristic of FT air (Figure 4a of the manuscript). Thus the June 20-26 time period was complex in terms of air movement, and thus it is not surprising that a global model will not capture this complex transport very well. Further evidence of the transport is given by the gridded frequency distribution of HYSPLIT back trajectories shown in the supplemental information, Figure 3. We disagree that this event should be removed from the FT chemistry discussion since it was observed under moderate to low WV concentrations indicating the site was situated in the FT and the observed RM/GEM ratio at MBO was close to the ideal -1 during the event, indicating FT photochemical processes. This unusual event had air masses that were in contact with the MBL several days back and we hypothesize that this is the primary reason that observed O₃ was relatively low. The revised manuscript has been improved to make this discussion clearer.

Comment #8: Section 3.4: In the last paragraph the authors claim that “.. this suggests that OH as an oxidant via the HgBr + OH could be more important in summer at the desert sites. . .”. This would, in the first approximation, require higher OH concentrations at desert sites than at MBO. In view of comparable O₃ mixing ratios but substantially lower WV concentrations at the desert sites than at MBO this seems unlikely. Another more general problem is that the HgBr + OH reaction appears from nowhere. A reader unfamiliar with the detailed reaction mechanisms would appreciate some beforehand information about

this reaction. He would also appreciate some explanation why this reaction was singled out from the dozens of other reactions.

Response to comment 8: The reviewer is correct that the conclusion that the HgBr + OH reaction is a bit of an over-reach. What we meant to say was that because we have found two distinct patterns when looking at RM/GEM ratios in the observations and in the model when the model chemistry was varied between the Br and OH-O₃ mechanisms, this suggests that there are some future experiments that could be tried. One is to look for the importance of Br and OH chemistry working simultaneously via the HgBr + OH pathway. If it is found that OH chemistry is more important at the desert sites, then the question that reviewer raises is important, why would this be so if OH concentrations are almost certainly lower due to lower WV mixing ratios? In the revised manuscript we have scaled back on our conclusions. We've noted the results when comparing the model with Br and OH-O₃ chemistry, and made suggestions for future research. The overall conclusion we make is that these results suggest there is not a single atmospheric oxidant for GEM and hence multiple chemical forms of Hg in air.

Comment #9: Page 22776, line 1: "Thus we run GEOS-Chem with the OH + O₃ kinetics to see where the Br mechanism might be deficient." This would work only if the OH + O₃ mechanism were without any deficiencies. Since the OH/O₃ mechanism has its own deficiencies (some of them are mentioned in the introduction), the authors check the deficiencies of one reaction scheme using another deficient reaction scheme. I think that this is the general problem of this and similar studies. An unbiased wording is needed.

Response to comment 9: The reviewer correctly points out a limitation in our study. We are comparing observations with known deficiencies, with model runs using chemical oxidation schemes that are far from proven. To mitigate this problem, we have attempted to present the data in an unbiased way and use the results more to inform future studies as opposed to prove hypotheses (beyond the obvious one that RM is produced in dry FT air). This kind of model-measurement comparison had not been done before for multiple surface site in the FT simultaneously and thus we felt even with the ambiguities of the results, it represents a step forward in our understanding of mercury science. That being said, we have removed the word "deficiencies" from our revised manuscript and phrased our intentions as: We will compare results from a simulation with Br chemistry versus one with OH and ozone chemistry. While both oxidation mechanisms, and possibly others, may operate together in the real atmosphere, these idealized simulations enable us to explore the constraints that observations place on the atmospheric chemistry of mercury.

Comment #10: Page 22776, line 28: "on average 2.8 ± 2.6 than the mean observed RM/GEM ratio" – average and standard deviation should always be accompanied by number of observations. No statistical tests for the significance of differences to zero or collection inefficiency estimate by Gustin et al. (2013) can be made when one of the three numbers (average, standard deviation, number of observations) is missing. This applies even more to the data in Table S1.

Response to comment 10: Values of N have been included in Table S1. N = 16 monthly means for the statistic 2.8 ± 2.6 .

Comment #11: In Section 4 the authors conclude that "this indicates that OH as an oxidant via the HgBr + OH pathway could be more important. . .". Does this reaction belong to the Br or OH/O₃ reaction scheme?

Response to comment 11: We have scaled back the assertion that $\text{HgBr} + \text{OH}$ pathway is occurring. This belongs in the Br reaction scheme according to Holmes et al. (2010).

Comment #12: Table S2: The units of RM/GEM, RM/O₃, and RM/WV slopes and intercepts should be given.

Response to comment 12: The units for the slopes have been included in the revised manuscript, supplemental information.

Reviewer #2

Comment #1: The authors extracted a subset of high elevation GEM and RM data of free tropospheric origin using water vapor mixing ratios. They then compared these data with GEOS-Chem simulated GEM and RM concentrations that used Br and OH-O₃ oxidation mechanisms. Examining high elevation data is interesting. However, I am afraid that the approach is problematic. Why should these two mechanisms exclude each other? They may very well work together in atmosphere Hg cycling. Several studies suggested that GEM+O₃/OH reactions are not realistic in gas-phase and lab studies suggested that they occurred in solid phase. In the atmosphere these reactions may still occur, but just not necessarily in gas phase; possibly they are in liquid and solid phase. However, the O₃/OH oxidation scheme in Selin et al. (2007) did not include heterogeneous chemistry, and those reactions are all gas-phase, which is apparently a seriously deficient mechanism. Holmes et al. (2010) published their work at the time to underscore the potential importance of GEM oxidation by Br; I doubt their intention was that Br was the only oxidant for GEM. In that scheme, my impression is that oxidation by O₃ and OH was not included. Their Br concentrations in the troposphere, excluding the marine boundary layer, were archived monthly averaged data, while in the MBL a value of 1 pptv was uniformly applied for BrO concentrations. The authors showed that RM/GEM ratios using the Br-oxidation scheme fall into two patterns, with a higher slope for the desert sites, and a lower slope for MBO and SPL, the mountain top sites, while the RM/GEM ratios using the OH-O₃-oxidation scheme from all sites generally fall along one line. To understand why the Br-mechanism caused that pattern, one might want to examine the distribution of Br concentrations for the desert sites vs. MBO and SPL, whereas O₃ and OH concentrations are orders of magnitude larger than GEM and hence the heterogeneous distributions of O₃ and OH may not affect RM/GEM ratios more than the inherent deficiency in the chemistry of the oxidation mechanism did. It is not clear to me how the authors arrived at the speculation, from that comparison, that “OH as an oxidant via the $\text{HgBr} + \text{OH}$ pathway could be more important in the summer at desert sites”. In short, the authors were using one deficient mechanism to prove the other deficient one incapable of doing a good job.

Response to comment 1: The reviewer correctly points out that it is problematic to compare observations with model runs using two different idealized reaction schemes, determine which one compared better to the observations and simply conclude that that model scheme represents reality to the exclusion of the other scheme. This was not our intention. Current ability to model with GEOS-Chem involves picking either of the two chemical oxidation parameterizations, not a combination of the two, or some other set of reactions. In spite of this limitation, as we state in the revised manuscript, “we will compare results from a simulation with Br chemistry versus one with OH and ozone chemistry. While both oxidation mechanisms, and possibly others, may operate together in the real atmosphere, these idealized simulations enable us to explore the constraints that observations place on the atmospheric chemistry of mercury.”

We mistakenly arrived at the conclusion that the $\text{HgBr} + \text{OH}$ pathway could be more important in the summer at desert sites, when in fact this is merely a hypothesis that could be tested with future GEOS-Chem work, which will investigate the effect of Br^- and O_3^- initiated gas-phase oxidation occurring simultaneously in the atmosphere, as well as aqueous and heterogeneous reactions. We have scaled back our main conclusion in the revised manuscript to state that since the Br^- and OH-O_3 chemistries compared differently to the observations, it is likely that there is not one single global oxidant for Hg^0 and as the reviewer points out, further work must be done before model-observation comparisons can yield insights into what the likely oxidation mechanisms are.

Comment #2: I wouldn't be surprised to see better correlation between modeled and measured RM concentrations using the OH-O_3 oxidation mechanism than that using the Br oxidation mechanism, because at least the chemistry was consistently produced in one model using the former mechanism, whereas the latter was apparently not, considering it used archived Br data from a different model. In another word, the better correlation possibly reflected the dependence of Hg cycling and OH/O_3 chemistry on solar radiation more so than the model's capability of simulating the chemistry that produces RM.

Response to comment #2: Differences in solar radiation are handled online in the GEOS-Chem Hg simulation. As the model runs in realtime, diurnal cycles of Br and OH are imposed based on the local solar zenith angle.

Comment #3: What were the authors suggesting by stating "Modeled RM/GEM using either oxidation scheme was on average 2.8 ± 2.6 higher than the mean observed RM/GEM, a factor roughly in line with the estimate of collection inefficiency of the KCl -denuder"? Does this collection inefficiency indicate measured GOM concentrations bias low? If that's the case, how should these data be used for model-measurement comparison?

Response to comment 3: As stated in the paper, there are known deficiencies with the RM measurements using the Tekran. These likely have to do with poor collection efficiency using the KCl -coated denuder and quartz fiber filter, and interferences with WV and O_3 . This was one reason for selecting the sites that experience FT air for this paper, since it is thought that these measurements have the least amount of bias compared to other measurements reported in the BL where humidity is much higher. However, there is a possibility that the Tekran RM measurements in this paper are biased low based on the results of Gustin et al. (2013) and Huang et al. (2013). We reported modeled RM/GEM using either oxidation scheme was 2.8x higher than observed RM/GEM and is in line with a factor of 3 that has been used recently to scale up GOM concentrations used in a dry deposition model. We felt this observation should be reported, even though our factor of 2.8 could be a result of other effects, like the model consistently sampling at a higher level in the troposphere, or wrong chemical mechanisms. It is true that using measurements with known deficiencies and comparing to model data generated with reaction schemes that may or may not represent the real atmosphere is a tricky proposition. The intent of the paper was to compare the observations with the model, run with two different set of assumptions, to see if any patterns would emerge that would point toward areas of future research both on improving the measurements and the models. We have revised the manuscript to make this objective clearer.

Comment #4: Also, I am not sure that using the 75th percentile water vapor mixing ratios as criteria to extract the FT data works for the two desert sites, because it is already very dry at those two locations.

Response to comment 4: See response to comment #4 from Reviewer 1.