Response to Comments by Anonymous Referee #4

The paper is well written and as the title suggests relates to an inter-comparison of 3 passive samplers for gaseous mercury in ambient air. Throughout the paper the analyte is referred to as gaseous mercury which by definition would normally include both elemental Hg (GEM) and reactive gaseous Hg species (RGM). There are no validation studies to my knowledge that report that passive samplers will determine RGM and therefore it would be more appropriate to refer the analyte as GEM and also provide additional discussion on this point.

The choice of the somewhat ambiguous term "gaseous mercury" (as opposed to "total gaseous mercury (TGM)" or "gaseous elemental mercury (GEM)") is deliberate, because not for all of the three samplers has the nature of the sampled mercury species been unequivocally established. While we agree with the reviewer that the likelihood of RGM to pass unhindered through the diffusive barriers of any of the three PASs used in this study is very small and therefore GEM is the most likely mercury species to be sampled, only for the *Mer*PAS® is there experimental evidence for this to be the case. Specifically, this paragraph is taken from the supporting information file of Si et al. (2020), which describes the use of the sampler developed by the University of Toronto, on which the *Mer*PAS® is based:

> "Previous publications presenting results using this PAS have referred to the analyte as gaseous Hg, but hypothesized that reactive gaseous oxidized mercury (GOM) is scavenged by the diffusive barrier (McLagan et al., 2016, McLagan et al., 2018a, McLagan et al., 2018b). Recent evidence has confirmed that GEM is indeed the mercury species sampled by the PAS. In an unpublished study (Stupple et al. 2019), two Tekran 2537/1130 speciation systems were run simultaneously under elevated GOM conditions during atmospheric mercury depletion events (AMDEs) at Alert in the Canadian High Arctic. After replacing the inlet of one system with the Radiello® diffusive barrier (no sorbent inside), the measured GOM dropped to near zero, while the other system continued to measure high GOM concentrations. This shows that GOM cannot pass through the diffusive barrier. Szponar et al. (2020) set up PASs with and without the Radiello® diffusive barrier during AMDEs at Alert. Isotopic analysis showed Hg isotope signatures distinctive of GOM in samples collected without the diffusive barrier, but not in samples collected with the diffusive barrier. Therefore, we are confident that GEM is an accurate description of the target analyte."

The use of passive samplers has limited value to fully understanding the cycling of Hg in the atmosphere as they only provide an average concentration over the deployment period. Generally, the legislation for ambient air only considers annual averages as representative of long-term exposure. In this respect, passive samplers add additional value to monitoring networks especially if they are deployed in remote locations.

First, we note that the minimum deployment period of a PAS could be quite short if atmospheric GEM concentrations are elevated. In areas affected by large GEM sources, resolution on the time scale of an hour or even shorter is feasible (Snow et al., 2021b). More importantly, as already stated in response to the comment by reviewer #1, we believe there is considerable merit to passive air sampling for GEM that goes well beyond long-term monitoring in remote locations. No single sampling/measurement approach is sufficient to fully understand the cycling of Hg. We note that neither automated ambient air monitoring by an instrument such as the Tekran 2537 can provide a full understanding of the cycling of Hg in the atmosphere as it can only provide concentrations at a very limited number of sampling sites. The key is that passive and active sampling are highly complementary and in combination are far more likely to get us the full understanding that we seek.

There is no discussion in the paper regarding manual sampling methods that use gold traps with a vacuum pump which is more comparable to passive samplers than an automated ambient air monitor such as the Tekran 2537. Manually sampling does require a power supply to operate the vacuum pump but they can also be used with rechargeable battery packs which for several weeks and they also offer total gaseous mercury (TGM) and much larger sample volumes.

The reviewer is correct that this type of sampler is not discussed in the paper. However, the paper does not address the question of the merits of different types of atmospheric mercury sampling techniques. It compares the performance of three different passive air samplers. While there are some potential applications of PASs for GEM that could also be accomplished with "manual sampling methods that use gold traps with a vacuum pump", many other potential PAS applications would not be amenable to that type of sampling. In the future, it may be worthwhile to pursue a similar inter-comparison exercise involving PASs and the type of sampling approach advocated by the reviewer.

There are several aspects of the analytical performance evaluation of the passive samplers in this paper that are questionable and therefore need revising.

The detection limit calculations are based solely on the variability of the field blanks and currently ignore the actual blank values which are significant in relation to the mass of mercury collected on the passive samplers studied. Reporting detection limits less than the blank valve is questionable as it is not possible to quantify a mass of mercury less than the blank. The authors should recalculate the detection limits according the following IUPAC expression (LOD = Blank + 3σn-1). This is far more appropriate way to report LODs for techniques that employ pre-concentration.

There are different ways of calculating limits of detection and the procedure we have used is commonly applied, widely accepted and scientifically defensible. If field blanks are consistently at a certain level (with a small standard deviation) it has no impact on the ability of a method to detect an analyte, even if that field blank level may be high relative to the amount of the analyte added during sampling. It also does not imply that a mass of mercury less than the blank is being quantified. It is the mass of mercury in the field blank plus the mass of mercury added to the sorbent during sampling that is quantified.

It is well known that sampling rates for passive samplers are affected by temperature, pressure, humidity and wind speeds. In this study these parameters are ignored even- though the metrological conditions at the sampling location have been measured and are available.

The reviewer is correct that in the current evaluation of the three passive air samplers, we have chosen to use sampling rates that are assumed invariant with respect to meteorological conditions, such as temperature, pressure, relative humidity, and wind speed. As stated in response to comments by reviewers #1 and #3, we decided to not consider the influence of meteorological parameters since the dependence of the sampling rate on temperature and wind speed has so far only been quantified extensively for one of the samplers, i.e. the *Mer*PAS® (McLagan et al., 2018a). However, as we reported on line 539, we observed a slight change of the SR.

The automated Tekran analyzers are based on a standardized volume measurement and the passive samplers are not. If the authors ignore the metrological conditions then the comparison is not valid.

This is not entirely correct. A Tekran analyzer does not measure the volume of air being sampled, but measures the mass of the air being sampled. In order to calculate volumetric air concentrations, that mass of air is converted into a volume of air applying the density of air at 0 °C and 1 atm. In a passive sampler, the sampled air volume is obtained from a sampling rate, which itself is the result of a calibration. What conditions apply to that volume, therefore, depend on the active sampling method that was used during the calibration. As in all cases a Tekran

analyzer was used during calibration of the PAS for Hg, the same conditions apply and the comparison is valid.

At both sites the metrological conditions were highly variable and in addition to this no pressure correction has been made for barometric pressure and elevation on each site. The impact of metrological conditions on sampling rates is discussed later on in the paper as an explanation for the higher variability at the Canadian site so why was the volumetric correction not applied? I would prefer that all passive sampler results are reported using a standardized volume based on the average metrological conditions on site for the deployment period.

As already indicated in the response to reviewer #3, one reason for assuming a constant sampling rate in our study is that the dependence of the sampling rate on temperature, wind speed, and relative humidity is not known for the IVL-PAS and CNR-PAS. It has only been quantified for the PAS developed by the University of Toronto on which the *Mer*PAS® is based (McLagan et al., 2018a). We have chosen to use invariant sampling rates in our comparison, and accordingly the assessment of bias and accuracy (e.g. Table 2, Figure 3) applies to a situation where a concentration obtained with a PAS while applying a constant sampling rate is compared with a concentration obtained with a Tekran.

It would also be good to know how the uncertainty of the sampling rate is affected by the metrological conditions. Whilst the overall correction might be small it is still important when performing and report the comparison. Humidity is unlikely to have an impact. Wind speed could have an impact but since the samplers are housed in a shelter then one could argue that this could also be ignored. The correction should therefore focus on temperature and pressure.

For the *Mer*PAS®, the dependence of the sampling rate on meteorological conditions is known quantitatively (McLagan et al., 2017, Snow et al., 2021a).

- Laboratory experiments at different temperatures established that the impact of temperature on the sampling rate is small (it changes by 0.0009 m^3/day for 1 K change in temperature) and can be fully explained by the effect of temperature on the diffusivity of elemental mercury.
- Laboratory experiments at different relative humidity found no effect on the sampling rate.
- Laboratory experiments at different wind speed established that the sampling rate at wind speeds > 1 m/s the dependence of the sampling rate is minor (it changes by 0.003 m^3/day for 1 m/s change in wind speed). At wind speeds below 1 m/s, the dependence becomes more pronounced. A sampler evaluation involving numerous field sites with a wide range of meteorological conditions also lends support to a slight wind speed dependence of the sampling rate (McLagan et al. 2018b).

Barometric pressure affects the PASs in two ways that cancel each other out. At lower pressure, the air is thinner and we might expect a sampling rate that is reduced proportionally to the lower pressure. On the other hand, mercury diffuses faster at lower pressure, which would increase the sampling rate, again to an extent that is proportional to the lower pressure. Lowering and increasing the sampling rate to the same extent means that the two effects cancel each other out and no adjustment should be necessary for a comparison with Tekran data, which are based on 1 atm of atmospheric pressure (i.e. sea level). We also note that Rende and Toronto are close to sea level in elevation and no notable impact of atmospheric pressure can be expected.

We could add a discussion of the dependence of sampling rates on meteorological conditions to a revised version, but are not convinced that this is necessary, as what is currently known about this topic has been described in detail in earlier publications.

The CNR-PAS has a more recent history compared to the other samplers, and as thoroughly discussed, the influence of meteorological conditions on the sampling rate has not yet been fully investigated. Recently a paper dealing with the effect of meteorological conditions has been published (Avossa et al., 2020). Besides, although the results of the presented study are promising (in line 539 we reported "A promising result of this study is that the SRs of the CNR-PAS at the two locations are more similar than for the other two PASs, which may hint at an SR that has a relatively small dependence on meteorological factors…"), further tests aimed at calibrating the sampler under a wide variety of meteorological circumstances are still in progress.

For the Accuracy assessment was performed by comparing the result the Tekran analyzers operating on each site. The paper explains that two Tekran instruments were operated but it seems that results from only one analyzer was used for the comparison? Surely, the comparison should be done on the average of the two Tekran results? If there are periods of known downtime for one of the two analyzers or technical justifications for ignoring a period of time then that is understandable but to ignore one analyzer and assume that the second analyzer is accurate does not seem correct.

As mentioned both in the manuscript and in the response to reviewer #1, the Tekran system labelled "0075" "experienced an 8 % shift in the mass-flow meter calibration during the study" and "since it was not possible to determine when the shift occurred" (line 268), it is not a matter resolved by identifying instrument downtime or separating data from before or after the shift occurred.

It seems that the authors have adopted a strategy of running analyzers in parallel but only using the second analyzer as reassurance them that the other analyzer was performing well. It would be nice to see the trend for both analyzers especially when so called spike events are experienced.

The "0075" unit did not meet QA/QC criteria and we were unable to correct the data as addressed in previous comments. Therefore, it would be misleading to present the data as a time series alongside the "5037" unit results which met all QA/QC criteria. In this case, the redundancy of having two systems running in parallel allowed us to have at least one Tekran system generating data suitable for comparison with the PAS, and allowed us to ascertain confidence in the active measurement.

Other concerns relate to the online analyzer reporting GEM even though the description of the analyzer suggests that TGM was determined because sampling was performed directly onto gold traps?

As indicated in a response to reviewer #1, we refer to the analyte in the manuscript as gaseous mercury, i.e. we avoid the nomenclature of TGM and GEM. We compare what is measured by a Tekran system with what is measured by the PASs. The PASs were calibrated with the same type of Tekran system, so the sampling rates are clearly appropriate for that analyte. In a revised version, we can make sure to not use the term "GEM", but consistently refer to gaseous mercury throughout.

The authors need to focus on what they are actually measuring with each technique applied. As mentioned above the Tekran measurements are based on a standardized volume whereas the passive samplers are not. It would be acceptable in the paper to discuss this in more detail and refer to comparability rather than accuracy. Having a passive sampler offering good agreement to the automated measurement is a very reassuring but to define the accuracy on this basis is not acceptable.

We believe we have addressed these points already. The fact of the matter is that the concentrations obtained by Tekran and PASs are comparable, if only because the PASs' sampling rate have been previously calibrated with the help of Tekran systems and therefore do apply to the same conditions. We also state explicitly in the paper (line 381) that "Tekran values were considered as a benchmark for pragmatic reasons, knowing full well that this measurement itself may provide biased results".

Overall this is very interesting study and the publication will be interest to many readers conducting mercury research in the field of atmospheric measurements. If passive samplers are deployed at more sites in the future then it is important to standardize the sample rate volumes especially when performing comparisons with automatic measurement systems.

We appreciate the reviewer's endorsement of the merits of our study.

References

Avossa, J., F. De Cesare, P. Papa, E. Zampetti, A. Bearzotti, M. Marelli, N. Pirrone, A. Macagnano. Characteristics and performances of a nanostructured material for passive samplers of gaseous Hg. *Sensors,* **2020**, *20*, 6021.

McLagan, D.S., M.E.E. Mazur, C.P.J. Mitchell, F. Wania. Passive air sampling of gaseous elemental mercury: A critical review. *Atmos. Chem. Phys.* **2016**, *16*, 3061–3076.

McLagan, D. S., C. P. J. Mitchell, H. Huang, B. Abdul Hussain, Y. D. Lei, F. Wania. The effects of meteorological parameters and diffusive barrier reuse on the sampling rate of a passive air sampler for gaseous mercury. *Atmos. Meas. Tech.* **2017**, 10, 3651-3660.

McLagan, D.S., B. Abdul Hussain, H. Huang, Y.D. Lei, F. Wania, C.P.J. Mitchell. Identifying and evaluating urban mercury emission sources through passive sampler-based mapping of atmospheric concentrations. *Environ. Res. Lett.* **2018a**, *13*, 074008.

McLagan, D.S., C.P.J. Mitchell, A. Steffen, H. Hung, C. Shin, G.W. Stupple, M.L. Olson, W.T. Luke, P. Kelley, D. Howard, G.C. Edwards, P.F. Nelson, H. Xiao, G.-R. Sheu, A. Dreyer, H. Huang, B. Abdul Hussain, Y. D. Lei, I. Tavshunsky, F. Wania. Global evaluation and calibration of a passive air sampler for gaseous mercury. *Atmos. Chem. Phys.* **2018b**, *18*, 5905-5919.

Si, M., D. S. McLagan, A. Mazot, N. Szponar, B. Bergquist, Y. D. Lei, C. P.J. Mitchell, F. Wania. Measurement of atmospheric mercury over volcanic and fumarolic regions on the North Island of New Zealand using passive air samplers. *ACS Space Earth Chem.* **2020**, *4*, 2435-2443.

Snow, M.A., M. Feigis, Y. D. Lei, C. P. J. Mitchell, F. Wania. Development, characterization, and testing of a personal passive sampler for measuring inhalation exposure to gaseous elemental mercury. *Environ. Int.* **2021a**, *146*, 106264.

Snow, M.A., G. Darko, O. Gyamfi, E. Ansah, K. Breivik, C. Hoang, Y. D. Lei, F. Wania. Characterization of inhalation exposure to gaseous elemental mercury during artisanal gold mining and e-waste recycling through combined stationary and personal passive sampling. *Environ. Sci.: Processes Impacts* **2021**b, DOI: 10.1039/d0em00494d.

Stupple, G.W., D.S. McLagan, A. Steffen. In situ reactive gaseous mercury uptake on radiello diffusive barrier, cation exchange membrane and Teflon filter membrane during atmospheric mercury depletion events. *In Proceedings of the 14th International Conference on Mercury as a Global Pollutant*, Krakow, Poland, **2019**, 8-13 September.

Szponar, N., D. S. McLagan, R. Kaplan, C. P. J. Mitchell, F. Wania, A. Steffen, G. Stupple, B. Bergquist. Isotopic characterization of atmospheric gaseous elemental mercury by passive air sampling. *Environ. Sci. Technol.* **2020**, *54*, 10533-10543.