This paper presents the results of model sensitivity studies of anthropogenic emissions and oxidation chemistry on atmospheric concentrations and deposition of mercury over Europe. The model results of each study are evaluated against observations.

General comments:
- I recommend the paper for publication after the revisions suggested here, which are largely about adding more explanation of the sensitivity studies and results, including more connection of the results to the current state of knowledge in the field, and highlighting implications of the results.

- In general, the discussion can be hard to follow at times since only some of the results are presented in the figures but results from all sensitivity studies are discussed.

In the different sections only the figures regarding the sensitivity tests that were useful to compare with observations are shown. Otherwise the article would have been overly long.

- Every option for emissions inventories, chemistry schemes, etc. are presented as equally valid but it would be helpful to know what discussion has already been performed of these inventories/chemical mechanisms etc. to indicate the strengths and weaknesses of each and how this paper’s results fit in. I think it is misleading to present Br and OH/O3 oxidation mechanisms on equal footing, for example, but then the paper shows that Br oxidation alone best represents TGM results which adds further support to what previous studies have suggested.

Text added: “A summary of the simulations performed is shown in Table 1. Some of these tests have been studied for other regions (e.g. Travnikov et al. (2016) and Bieser et al. (2016)) while many other studies have investigated Hg oxidation by Ozone or Br (Hynes et al., 2009; Subir et al., 2011, 2012; Weiss-Penzias et al., 2014).”
So in this paper we are not trying to make a direct comparison between the two mechanisms, but simply an analysis of how much the atmospheric Hg cycle may change considering a single Hg oxidant. The above text was added in section 2.3.

- I think the paper would benefit from more clearly outlining the purpose of each of the sensitivity studies. It would also be beneficial to have the explanation of most sensitivity studies be more centralized – since some of the sensitivity studies are untrue hypotheticals (e.g., NOCHEM and NOANT) while others are testing different hypothesized physical processes (e.g., BRCHEM1 vs. BRCHEM2), but both types are valuable. There is some of this scattered throughout the results section, but I think it can be more explicit; e.g., ‘the ANTSPEC experiment, which assumes all anthropogenic emissions are as GEM, would represent a lower bound on deposition from local anthropogenic sources and an upper bound on long-range transport of anthropogenic emissions because GEM has a much longer lifetime against deposition than RGM/PBM’ or something to that effect.

Very useful suggestions. Section 2.3 was expanded adding this information (bullets of different experiments).

- Some background information is missing – e.g., the lifetime of GEM vs. RGM/PBM is not explained in the paper, or the sinks of GEM vs. RGM/PBM (dry deposition vs. wet deposition
affinities), making the results from the ANTSPEC simulations and others less easily understood. Some aspects of atmospheric mercury are missing from the discussion (e.g., reduction of Hg(II), uncertainty in present oxidized Hg measurement capabilities that would be relevant for the 2nd to last sentence of the paper).

This information has been added in the Introduction

Specific comments:
Section 2.1: There are a lot of details given here about the models but it is not clear how the model representations of Hg are different or how their differences would affect Hg results. It would be helpful to rewrite this a bit and make it more explicit whether/how the differences would affect Hg (e.g., if CMAQ is offline while WRF is online meteorology, the transport of Hg will be different to some degree between the models.)

A more extensive models description was added in Section 2.1

Page 4 line 2: it would be helpful to explain briefly here whether there are multiple heights in the EDGAR inventory, as it is written it is unclear what the difference is and this comes up later in the results section.

Text added: “They also have different emission height distributions: AMAP uses three height classes (0-50, 50-150 and above 150 m) whereas EDGAR ranges into six classes (distributed between 0 and 800 metres, listed according with SNAP (Selected Nomenclature for Air Pollution) categories as used in De Simone et al. (2016)). The differences in the geographical distributions are shown in Fig. 2. The EDGAR anthropogenic ranges in six classes, distributed between 0 and 800 m, listed according with SNAP categories. This was added in Section 2.2.

Section 2.3: It is unclear from the way the paper is written what the base model oxidation chemistry schemes include –is it OH, O3 and Br oxidation at the same time in both CMAQ and WRF-Chem Hg? This should be made clear.

The sentence was modified to explain this point: “The BASE experiment refers to the model in the standard configuration, with AMAP anthropogenic emissions distributed according with the inventory guidelines.” has been changed to “The BASE experiment refers to the model in the standard configuration, with AMAP anthropogenic emissions and Hg oxidation driven only by O3 and OH for WRF/Chem-Hg and by O3, OH, H2O2 and Cl2 for CMAQ-Hg, as described in section 2.1.”

Page 4 Lines 23-25: some of this would be helpful to mention in the model description section 2.1.

Sections 2.1 and 3 were modified following this suggestion.

Table 3: add a total deposition column, just makes it easier to follow

Table 3 was modified following this suggestion.

Figure 3: It is hard to compare the results between CMAQ and WRF-Chem Hg because the same model set ups do not line up. I think it is worth reordering the individual panels to lining up the ANTSPEC and ANTSPEC_C vertically, or NOCHEM with NOCHEM_C below it, etc. to make more obvious the comparisons made in Section 3 of the paper.
We are not sure that it is convenient to change the order of the figures. The maps should be read horizontally, with the first column (BASE, ANTSPE and BASEc) as a reference for the other maps.

Figure 4: This is total dry deposition of Hg(II)+Hg(0), correct? Just making sure since wet deposition is of course just Hg(II). I think this is worth mentioning as it also helps explain the changes in dry vs. wet deposition seen in the sensitivity studies.

No, it is just wet deposition flux, not the sum. Only the total dry (figure 3) and total wet (figure 4) deposition contributions are reported in the figures, to underline the results of tests. Total dry and total wet deposition are obtained as the sum of gaseous species and particulate deposition.

Figure 5a: there are two legend labels with “BASE”. I assume the green square is BASE2? Also, I would explain that the order of the sites is by the magnitude of the mod/obs ratio of the BASE simulation as it is not immediately obvious.

The legend of figure 5 was modified following this suggestion.

Page 5 Lines 20-23. It is counterintuitive to me how an overestimate relative to observations for the ANTSPEC sensitivity study where all anthropogenic emissions of RGM & PBM are as GEM means long-range transport is less important than regional emissions. I think more explanation would be helpful – so when PBM and RGM are emitted normally as in the BASE simulation, the model is no longer overestimating the observations because the Hg(II) is deposited fast enough before it reaches the CZ03 site? I suppose part of this is defining what is local vs. “regional” vs. long-range more clearly.

First, the definition of Local and Long-Range follows that of Gencarelli et al., 2015: Local are the sources of Hg inside the domain (anthropogenic emissions and evasion from the sea surface), Long-Range sources of Hg come from boundary condition (external to domain). About the overestimate relative to observations for the ANTSPEC experiment, it concern also the GEM air concentrations (I was wrong not specify this), in fact the depositions in CZ03 station are lower in the ANTSPEC than the BASE experiment. Section 3.1 was modified following this suggestion.

Page 5 last paragraph (beginning line 29): do you have any hypotheses why TGM at ES08 is so low relative to the models?

From the modelling point of view of this study no clues have emerged to explain these low concentrations. The causes are probably attributable to the measurement technique or some unusual combination of local phenomena.

Page 6 Lines 1-2: The discussion of OHCHEM and NOANT experiments comes as a surprise as it is not in Figures 5, 6, or 7; more connection to the rest of the section is needed.

In order to include OHCHEM and NOANT to the rest of discussion, the results of these experiment are reported in figure 5. It is not surprising that this sensitivity test yields very low concentrations in comparison with total deposition and overestimation with air concentrations.

Page 7 Lines 15-16: it is also not immediately obvious why dry deposition decreases so much more than wet deposition, since GEM is not wet deposited. Is it because even though there is higher GEMair in the ANTSPEC experiment, it is dry deposited so much more slowly than the RGM/PBM species, or is it something about how RGM and PBM contribute different amounts to wet vs. dry dep of the oxidized species? Setting up the background on this in the intro or methods would be helpful.
This explanation has been added to section 3.2:
In the ANTSPEC experiment the deposition decreases in comparison to BASE, in particular the dry decreases more than the wet. % RGM and PBM deposit more rapidly than GEM and so deposit in proximity to their emission sources where the air concentrations are higher. Clearly dry deposition can occur at any time while wet deposition requires precipitation. With all $\text{Hg}$ emissions releases treated as GEM in ANTSPEC the dry deposition decreases most as a result of the lack of direct emissions of RGM and PBM.

Page 7 lines 21-23: given that a no anthropogenic emissions scenario is currently untrue, I think something more insightful can be said about the results – e.g., something about how anthropogenic emissions contribute to $2/3$ of total deposition (not counting the fact that “natural” emissions from soil/ocean as they are tuned in models also implicitly are impacted by legacy anthropogenic sources), or how a hypothetical policy scenario of shutting off emissions could have huge local benefits?

The final part of section 3.2 was modified according to this suggestion and the results of Pacyna et al., 2016, that describes the results of modelling studies using GMOS project scenarios to assess $\text{Hg}$ concentration and deposition fields, for present (2013) and future anthropogenic (2035) $\text{Hg}$ emissions.

Section 3.3 paragraph 2 (pg 7 lines 30-32): I would emphasize that this shows a significant proportion (exact percent varying on the model) of total $\text{Hg}$ deposition to ecosystems is coming from the oxidation of GEM which can be transported from far distances as opposed to the $\text{Hg(II)}$ locally emitted. This is an interesting result with policy implications and could be highlighted more.

Section 3.3 and the conclusion have been modified according with this very useful suggestion.

Page 8 lines 3-4: “A number of studies have shown the importance of O3, and the OH radical, and also reactive halogen compounds. . .”: I understand that there have been review papers discussing the intricacies of this and you don’t want to repeat that here, but it is overly simple to group all three oxidants together and not mention that studies have found that the homogeneous gas-phase oxidation of $\text{Hg(0)}$ by O3 and OH are thermodynamically and/or kinetically impossible (e.g., Hynes et al. (2009), Goodsite et al. (2004), Calvert and Lindberg, 2005). I think it is still interesting to compare the results from the three species, but it needs to be introduced with a bit more nuance. Moreover, as Theodore Dibble posted in his comment, there are additional $\text{HgBr}+X$ second-step oxidation reactions that can greatly increase the total $\text{Hg(II)}$ production and deposition through the Br-initiated pathway. Somewhere in the paper there should be a discussion of how this would affect the results presented – e.g., $\text{Hg}$ deposition in the BRCHEM1 and 2 sensitivity simulations would be increased and TGM would be decreased.

The section 3.3 was modified according with this suggestion.

Page 9 Lines 1-8: it is hard to understand the differences in the two Br concentration fields from the description given – would it be possible to show (perhaps in the supplemental) a difference plot of the Br concentrations over the Europe domain? (e.g., zonal mean latitude on x-axis vs. altitude on y-axis or something like that). Most GEM oxidation is not occurring in the PBL but in the free and upper troposphere (because of Br distribution and the temperature dependence of the oxidation reactions), so it is not surprising that the huge differences in Br in the PBL between the two Br fields doesn’t impact on $\text{Hg(II)}$ deposition; I am more interested in the differences at higher altitudes.
To supplement the information about Bromine concentrations the vertical longitudinal profiles of the annual mean concentrations were added in section 3.3. The main differences between the two inputs are in the first 3 km and in the range 12-15 km, as shown in the new version of section 3.3.

Page 9 Lines 13-16: I think it is essential to connect this to available observations of Br. The Shah et al. study tripled bromine concentrations of the GEOS-Chem Parrella et al. 2012 model which was consistent with observations of BrO during the NOMADSS field campaign (Gratz et al., 2015). Parrella et al. (2012) showed previously that BrO was underestimated in GEOS-Chem by 30% in the global mean against satellite observations. So it is not just a model exercise but shows that higher Br in GEOS-Chem (closer to reality but not quite there) is also closer to reality for Hg model results. But in addition – since anthropogenic Hg(II) emissions have been turned to GEM in reality there would be more deposition with just Hg+Br oxidation and anthropogenic emissions of RGM/PBM turned on. This makes me curious about the Hg(II) reduction mechanisms in CMAQ vs. WRF-Chem Hg – if this was included, was it treated the same in all sensitivity simulations? I think reduction needs to be discussed somewhere in the paper.

The last part of section 3.3 was modified in accordance with this useful suggestion.

additional references cited here:


