

The authors discussed the comments in the response letter. However, only a few of the comments are addressed in the manuscript. I recommend the authors go through all the comments again and at least include most of the comments in the revised manuscript. In addition, I still have some comments that are not well discussed in the response letter. I think at least a major revision is needed.

Other specific comments:

1 Before inversion, the prior emissions were pre-optimized to fit the global mean MCF mole fractions. The authors argue that the pre-optimized MCF emissions can reduce computational costs. But the inversion as shown by equation(1) is to estimate the emissions and OH by combining the information from both bottom-up estimated prior emission inventories and the observations, as well as their errors. The pre-optimize erase the information of prior emission inventories, and only keep the information of observations. I don't think this is the right way to do an inversion. Besides, the question is not answered: "Will the pre-optimization reduce the OH variation estimated by the inversions since the MCF emissions already fit the observations?" In addition, from the author's response to Fig.4, I feel the pre-optimization is somewhat arbitrary, which makes the inversion lost the prior information.

2 The authors explain why the small correlation coefficient between OH and MCF variations can reflect the OH and MCF are independently derived. But I think a better indicator should be comparing optimized minus prior MCF emissions and optimized minus prior OH. We can see that in Fig.3, the MCF emissions estimated by REF inversion is much higher than prior around 2013, which is corresponding to the large positive OH anomaly around 2012-2013. This may indicate that the inversion system cannot separate the OH and MCF variations.

3 For the convergence problems. From the author's discussion, the 10-year inversions can reach convergence since they require less time per iteration. One problem is that the 10 years inversions are for 1998-2008 when the MCF emissions are higher than 2009-2018 and the corresponding errors are much lower than 2009-2018. If the 10 years inversions focus on 2009-2018, it will be hard to say if the 10-year inversion coverage to similar OH variations since the uncertainties in MCF emissions (reach 200%) are much larger OH during 2009-2018. So I don't think the 10-year inversion for 1998-2008 can prove the robustness of the 20-year inversions, as the author mentioned when discussed my last comment.

4 "L352: Why further convergence will result in less realistic OH variations?"

The authors answer this question by adding Figure S8 (but show nothing in the manuscript) which showed that the inversions are overfitting. Is this because the inversions use too small observational error?

5 "L372: "Firstly...we generally identified similar tendencies in each." Figure 1 has

shown the variations of OH estimated by three inversions are quite different.”

The authors answer this by showing the 10 years inversions are similar. But as aforementioned, the three 10-year inversions are similar may not prove the inversions for 2009-2018 can also reach similar results. Here the only thing we see in the main text is that the three 20-year inversions are quite different.

6 “L374-379: Here the manuscript tried to prove the robustness of the OH interannual by an additional inversion and a forward simulation. But the details of the two experiments and the results are not given. I suggest include some details in the supplements. E.g. how the one global scaling factors compare with the REF, POP, and TM5OH? Is the forward simulation use the prior or optimized MCF emissions? I cannot understand the logic here, can you clarify why the two experiments can indicate the robustness of the derived OH variations?”

It is still unclear how the authors conduct the two experiments. I think every model experiment established should be introduced in the manuscript or supplements. Since the two model experiments are not shown clearly, the role of the two experiments is certainly unclear.