L349: What are the criteria for stopping the iteration? Are the 3 inversion experiments reach a similar value of cost function J (and the gradient of the cost function) in the last iteration?

The inversions were stopped when additional iterations did not result in further modification of the solution, e.g. the OH variations. Because of the reasons outlined in this section, this did not necessarily mean that the cost function gradient was reduced by a factor that we pre-defined (typically 1000). Rather, we consider the stagnation of the inversion indicative of convergence problems.

The REF and TM5OH 20-year inversions converge to a similar value for the cost function, while the POP inversion ends up somewhat higher. However, this is just one number: we consider an analysis of the error statistics of the solution more insightful, such as in Figure 6.

The norm of the cost function gradient is typically reduced by a factor ~100 in all three inversions, but this value strongly oscillates between subsequent iterations.

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

The 10-year inversions find large adjustments in the latitudinal distribution of OH that fall significantly outside the spread of a wide range of literature estimates (see the new Figure S8). Therefore, we did not consider it worth the computational cost to continue the 20-year inversions only to end up closer to the 10-year inversions, especially since we consider some aspects of the 10-year inverse solution physically unrealistic.

L372: “Firstly...we generally identified similar tendencies in each.” Figure 1 has shown the variations of OH estimated by three inversions are quite different.

Our argument is that this difference comes from convergence issues and not from a difference in the converged solution. The evidence for this is that the 10-year inversions do convergence to the same OH variations, which makes it very likely that the 20-year inversions would also, given infinite computational time. Moreover, in Figure 1 it can also be seen that there are similarities between the three inversions in the timing and sign of derived OH variations, only these OH variations are most pronounced in the REF inversion. More precisely, a positive OH anomaly in the REF solution never coincides with a negative anomaly in the TM5OH and POP solutions, and vice versa.

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?

The point we want to make here is that the interannual variations in global mean [OH] (or oxidation) give an improvement relative to interannually repeating OH that is independent of the large latitudinal OH adjustments we find. I.e., even if the latitudinal biases were less strong (for example if we adjust the ocean flux) we would likely still find the same OH variations.

The simplest way that we have shown this is by only including the global OH variations (as in Figure 1) in a forward simulation, and not the latitudinal OH adjustments. This forward simulation reproduces the observations better than a forward simulation with no interannual OH variations. We find this result even if we use prior MCF emissions in both simulations (although, obviously, if we also include optimized MCF emission variations, the result improves further).

The second test that we performed in an early stage of the study is to only optimize one global scaling factor for OH. In this case, we still derived similar OH variations. We choose not to present the inversions with one OH scaling factor, because in these test inversions several other input variables were different from the inversions presented in the main manuscript (e.g. the prior emissions). Therefore, a fair presentation of these results would require an extensive and complex explanation. We are afraid that the storyline of the paper would become even more convoluted.

We only mention these two experiments to underline the (minor) point that the derived global OH variations are independent of the latitudinal OH adjustments. The general robustness of the derived OH variations is best shown by the ten-year inversions, which we do discuss extensively.