

This work attempts to attribute the sources contributing to the atmospheric CH<sub>4</sub> mixing ratio and their trends in China using the GEOS-Chem model simulations driven by two commonly used global anthropogenic emission inventories. It uses in-situ and satellite observations of CH<sub>4</sub> mixing ratios to explain the model results. Study also performs sensitivity test with OH to reproduce observed CH<sub>4</sub> mixing ratios and trends over China.

The discussion on the differences between the model results and observations is not sufficient. Authors can address some missing pieces of information or address some limitations in the results shown. I recommend the manuscript send back for the major revisions with following major/minor comments:

Major comments:

The study claims “model simulation using the CEDS inventory and interannually varying OH levels can best reproduce observed CH<sub>4</sub> mixing ratios and trends over China”. I don’t agree to a certain extent.

First of all, EDGAR v4.3.2 provides global emission estimates, at source-sector level, for the historic period from 1970 until 2012. How did the author estimate the EDGAR emissions beyond 2012? It appears the emissions are extrapolated (?) till 2018 for this study. Similarly, in case of CEDS inventory the emission estimates are during 1970-2014. How does the emissions are calculated beyond 2014 in this case too?

In Figure 6, over ‘DSI’ and ‘LLN’, simulations from both inventories are comparable at least till the year 2016 (it appears that, trends are affected by later years simulations for GCE). Over ‘SDZ’, EDGAR performs better than CEDS, however, over ‘WLD’, CEDS is better than EDGAR. Overall, these results are not very conclusive to say CEDS is better.

In Figure 7-8, the trend correlations for model simulations using EDGAR and CEDS with GOSAT are not significantly different.

In Figure 9, it appears for HIPPO observations, both simulations (performed using EDGAR and CEDS emission inventory) are within observed standard deviation. But for ATOM observations, CEDS inventory performs better than EDGAR. One reason to me is EDGAR extrapolated(?) emissions are used for the model simulations comparison with ATOM observations, whereas, in case of HIPPO observations actual EDGAR emission estimates are used. So, this Figure is also not very conclusive to say CEDS inventory is better, moreover, almost all the observations from HIPPO and ATOMS are over Pacific and American continent.

Figure 10, mixes both aircraft observations, which is not correct in my opinion. This figure is confusing.

Another issue is the source attribution of CH<sub>4</sub>. The attribution of CH<sub>4</sub> sources with tagged tracer needs more evidences. The source contribution should be provided along with confidence interval. Is there any relevant study to support this analysis for CH<sub>4</sub>?

Minor comments:

Some places in the manuscript authors use 'Fig.' and somewhere 'Figure'. Please use uniform convention.

Line 55; please add a reference after “a lifetime of 9.14 ( $\pm 10\%$ ) years”

Fig2: How do you define the regions for tagged CH<sub>4</sub> tracer simulations?

In Figure 6-10, please mention the model configure and OH field configuration used for simulations in the caption for better clarity.

Fig10: Legends needs to be adjusted properly.