

We greatly appreciate your thorough review and helpful comments and suggestions. A point-by-point reply to the comments is as follows.

General comments

First, the manuscript should undergo extensive language editing. Although I am not a native speaker of English, I can notice that a lot of sentences in this manuscript, mostly in the Introduction section, are composed of too many clauses and are often hard to understand. Second, section ‘4 Historical changes’ can be improved, in both contents and structure (see below for detail). Third, the use of CMIP6 data in comparing with FireMIP model simulations sounds like a lit of circular argument to me, since results from 6 FireMIP models were used in the creation of CMIP6 reconstruction. I believe this paper will be an important contribution to the fire community once these issues are adequately addressed.

- Reply:** (1) The language has been edited extensively to improve the readability.
- (2) Contents and structure in Sec. 4 have been revised. Please see response to your specific comments below for details.
- (3) This study provides and analyzes simulation data from each of nine FireMIP models. Fire history in CMIP6 data is estimated using fire proxy data (charcoal records and visibility records) for North America, Europe, Equatorial Asia, and central Amazon, and only the median of the simulations from six FireMIP models in each grid cell for other regions. Fire proxy data are independent of FireMIP model simulations. Multi-model medians/means are sometimes used as benchmarks to compare with simulations of single models in Earth system research (e.g., Lawrence et al. 2007, Journal of Hydrometeorology), so we think it is appropriate to compare them although they are not entirely independent. For clarification, we have changed “FireMIP models” to “median of six FireMIP model simulations” when describing the sources of CMIP6 fire emissions in Table 5 (Table 4 in the old version).

Specific comments:

1. Complex or ambiguous sentences in the ‘Introduction’ (an incomplete list):
 - 1) L66-69: This sentence seems too complex. The four ‘and’ and one ‘as well as’ used in this single sentence make it hard to be understood.

Reply: The sentence has been rephrased as “Our study provides an important dataset for the development of regional and global multi-source merged historical reconstructions, analyses of the historical changes of fire emissions and their uncertainties, and quantification of their role in the Earth system.”

- 2) L81-89: Similarly, this sentence is way too long. The last clause (regarding the ‘air

quality') should belong to a separate sentence.

Reply: According to your suggestion, we have divided the sentence into three as

“Second, by changing the atmospheric composition, fire emissions affect the global and regional radiation balance and climate (Ward et al., 2012; Tosca et al. 2013; Jiang et al., 2016; Grandey et al., 2016; McKendry et al., 2018; Hamilton et al., 2018; Thornhill et al., 2018). Third, fire emissions change the terrestrial nutrient and carbon cycles through altering the deposition of nutrients (e.g., nitrogen, phosphorus), surface ozone concentration, and meteorological conditions (e.g., diffuse radiation, temperature, precipitation) (Mahowald et al., 2008; Chen et al., 2010; McKendry et al., 2018; Yue and Unger, 2018). In addition, they degrade the air quality (Val Martin et al., 2015; Knorr et al., 2017), which poses a significant risk to human health...”

3) L93-94: The authors are too assertive in some claims and statements, in my opinion. For instance, in both cases of ‘fire emissions are estimated based on. . .’ and ‘Satellite based fire emission estimates are derived from. . .’, it may be better to use more modest expressions such as ‘are often estimated. . .’, or ‘are primarily derived from. . .’

Reply: We have revised sentences that are too assertive. For instance, the two sentences you mentioned have been changed to “are often estimated...” and “are primarily derived from...” as you suggested.

4) L98-99: ‘Data are available globally, but only cover the present-day period’. What ‘Data’ are you exactly talking about, (general) fire emission data, or satellite-based fire emission data? Please be more specific.

Reply: “data” has been changed to “Satellite-based fire emission estimates”.

5) L100-101: ‘and CO concentration trapped in. . .’. It is the CO who gets trapped, not the ‘concentration’.

Reply: “records of..., and CO concentration trapped in ice cores” has been changed to “ice-core records of..., and CO”

6) L104-108: Again, I have a problem in understanding this ‘complex’ sentence, partly C2 due to the 6 ‘and’/‘or’ appearances in the final clause.

Reply: The sentence has been divided into two as “The fire proxies can be used to reconstruct historical fire emissions on a local to global scale and for time periods of decades to millennia and beyond. However, fire proxies... ”

2. Section 4: Historical changes: Sections 4.1 and 4.2 are not well separated (even their titles are similar). The drivers of historical changes are discussed at the end of 4.1 and also in 4.2. Is it better to move all contents of drivers to section 4.2, and switch the section titles of 4 (Historical changes) and 4.1 (Historical changes and drivers)?

Reply: We agree with the reviewer.

We have moved the discussion about the content of drivers from Sec. 4.1 to Sec. 4.2, and changed the titles of Secs. 4, 4.1, and 4.2 to “Historical changes and drivers”, “Historical changes”, and “Drivers”, respectively.

3. L359-360: Any theoretical explanation on the lower amplitude of seasonality from JSBACH-SPITFIRE model?

Reply: We have added “likely caused by parameter setting in its fuel moisture functions (Table S9 in Rabin et al. 2016)”.

4. L440-441: Can you expand the explanation a little bit? i.e., how did ‘assuming no fires over croplands and setting high fuel bulk density for pastures’ lead to the sign change in LULCC response in JSBACH model?

Reply: As suggested, we have expanded the explanation to

“In JSBACH-SPITFIRE, as croplands and pastures expand over time, the assumption of no fire over croplands tends to decrease fire emissions, while the setting of high fuel bulk density for pastures tends to increase fire emissions due to increased fuel combusted per burned area, which together partly result in the shifted sign of response to LULCC around the 1940s.”.

5. Section 4.3: I like the discussions of drivers of global changes in section 4.2. But I would also like to see how these drivers play different roles on a regional scale.

Reply: We have added 14 figures in the supplementary material (Figs. S6-19) which are similar to Fig. 7 (global) but for 14 regions, to show the role of drivers on a regional scale.

Also, we have added a new paragraph to briefly describe them in Sec. 4.3 as
“The long-term changes of regional fire emissions and inter-model disagreement are mainly caused by simulated responses to LULCC and/or population density change for the 20th century (Fig. S6-19). Besides, climate change also plays an important role in North America, northern South America, Europe, northern Africa, boreal and central Asia, and Australia. FireMIP models generally simulate increased regional fire emissions with increased CO₂ concentration and negligible impacts due to changes in lightning frequency, similar to the responses of global fire emissions.”

6. Possible circular reasoning. According to the text in L303-308, CMIP6 estimates were calculated using different data sources (including 6 FireMIP model results). But the details of the reconstruction process were not given in the manuscript. How large do the FireMIP model results contribute to global emissions in CMIP6? Regardless of the amount of this fraction, some agreements between FireMIP and CMIP6 shown in Figures 6 and 9 are likely due to the use of the same data source. If you plot similar figures using data in North America + Europe + Equatorial Asia + central Amazon

(where no FireMIP information is used in CMIP6) only, the comparisons will be independent and maybe more convincing.

Reply: Please see the response to your general comment for the comparison between FireMIP simulations and CMIP6 estimates above.

We have revised the Fig. 9, which now provides a comparison between CMIP6/CMIP5 and simulations of FireMIP models in boreal North America, temperate North America, Europe, Equatorial Asia, NH South America, and SH South America. A brief description about them are in the revised Paras. 2 and 3 of Sec. 4.3.

In addition, Figs. 8-11 in van Marle et al. (2017, paper for CMIP6 fire emissions) already compared simulations of FireMIP models and their median with historical fire emission reconstructions based on charcoal records and visibility data (i.e. CMIP6 estimates) in four sub-regions of North America, Europe, and Equatorial Asia, and central Amazon.

7. L330: It will be interesting to see the combustion completeness ranges in FireMIP models other than LPJ-GUESS-GlobFIRM.

Reply: We have added combustion completeness ranges of all FireMIP models in Table 2, and have changed the sentence to "...than those used in other FireMIP models (Table 2) and the satellite-based GFED family (20–40% for stem and 40–60% for coarse woody debris) (van der Werf et al., 2017)."

8. L492: ‘fire and Earth science research communities’. Is fire science not a part of the Earth science?

Reply: Fire is a part of the Earth science. “fire and” has been removed.

9. Figure 1: ‘CRUNCEP atm.’ shown in this figure is not easy for readers who are not familiar with reanalysis data. This can be changed to ‘atmospheric forcing’ as being consistent with that in the main text.

Reply: “CRUNCEP atm.” has been changed to “atmospheric forcing” in Fig. 1.

10. Figure 3. The pattern shown in this figure is highly dependent on the spatial distribution of BC emissions. It will be good to see a map of inter-model std normalized with mean emissions.

Reply: We plotted the inter-model std normalized by mean emissions for grid cells where mean fire BC emissions were larger than $0.001 \text{ g BC m}^{-2} \text{ yr}^{-1}$. High values were located in regions with small mean emissions, which were in fact not important for the global fire emissions, e.g., arid regions, central rainforests. Thus, we decided to keep the inter-model std map in the manuscript.

11. Figure 7: The population density is shown in the figure as ‘control run -

sensitivity run’, which may cause a lot of confusion. In fact, I had a hard time understanding the meaning of ‘increasing population density’ (in L416) and ‘rising population density’ (in L421) at first, until I realized the use of this reverse scale in Figure 7. Is there any particular reason that you did not use ‘sensitivity run - control run’ instead?

Reply: Compared with ‘response to no population density change, no climate change,...’, we think ‘response to population density change (rising/increasing population density over the 20th century), climate change,...’ is more intuitive and helps better understand the simulated fire emission change shown in Fig. 6, so we used a reverse scale/’control run - sensitivity run’ in Fig. 7.

To help understand Fig. 7 and related text easier, we have briefly described the control and sensitivity runs in the caption of Fig. 7 and the 20th century change of driving forces used in FireMIP in both the caption of Fig. 7 and Sec. 4.3 .

Technical corrections:

1) L59: ‘most of the models’ to ‘most models’

Reply: Done

2) L116: Is it better to change ‘are applied to global change research’ to ‘have been widely used in global change research’?

Reply: Yes, changed accordingly

3) L142: In order to make it more specific, ‘Our study’ may be replaced with ‘This study’, or ‘The present study’, or ‘The study presented in this paper’, etc.

Reply: “Our study” has been changed to “This study”

4) L144: ‘the nine DGVMs’ to ‘nine DGVMs’

Reply: Done

5) L145: ‘The dataset provides the basis for’ to ‘This dataset provides a basis for’?

Reply: Done

6) L280: Why not spell out ‘CE’ for easier reading?

Reply: We have spelled CE out as “fire carbon emissions”.

7) L325-326: ‘whereas they are 1.5-4.2. . .for satellite-based products’. To be consistent with the previous clause, the range value should be in the singular form.

Reply: Changed “they are” to “it is”