

Response to reviewer 1

Response: We thank the reviewer for his constructive comments. Here we only list review comments that need to be addressed in the revised version (i.e. not comments that were answered positively).

Specific comments:

3. There is no such section as ‘Conclusions’ in the paper. Therefore, it is not clear whether this paper is aimed to reach some substantial conclusions. It seems that it is focused on “problem statement”. “There is a problem”, could be interpreted as a substantial conclusion if the problem is formulated clearly and supported by an analysis of the “state-of-the-art”. However, authors should do some effort in this direction. In the present version of the paper, I did not find something that could be considered as a “substantial conclusion”, although it is quite obvious for me that the paper may lead to substantial conclusions.

Response: The reviewer is correct that this paper is intended to be a statement of a problem, supported by an analysis of the state-of-the-art. Succinctly, we need to be able to predict how fire regimes will change in the future and in order to have confidence in these projections we need to evaluate how well different models perform, and to devise ways of improving model performance if this is necessary, and we have set up the FireMIP in order to make some progress with these goals. We are sorry that this message (or conclusion) did not come through clearly enough. We have changed the title of the final section from “the next steps” to “Conclusions and Next Steps” and expanded the text to strengthen this important point. Specifically, we have added a first paragraph to this section to emphasize that the goal of FireMIP is to demonstrate whether existing fire models are sufficiently mature to be used for projections as follows (Lines 468-473):

“Fire has profound impacts on many aspects of the Earth system. We therefore need to be able to predict how fire regimes will change in the future. Projections based on statistical relationships are not adequate for projections of longer-term changes in fire regimes because they neglect potential changes in the interactions between climate, vegetation and fire. While mechanistic modelling of the coupled vegetation-fire system should provide a way forward, it is still necessary to demonstrate that they are sufficiently mature to provide reliable projections. This is a major goal of the FireMIP project.”

The existing text for this section emphasized the different levels of complexity of existing fire models and the fact that we do not yet know what level of complexity is required to achieve robust results. We have preserved this paragraph, but have revised the final part of this paragraph to point out that another major goal of FireMIP is to establish the level of complexity required as follows (lines 488-492):

“FireMIP will address these issues by systematically evaluating the performance of models that use different approaches and have different levels of complexity in the treatment of processes, in order to establish whether there are aspects of simulating modern and/or future fire regimes that require complex models. Systematic evaluation will also help guide future development of individual models and potentially the further development of vegetation-fire models in general.”

We have used the opportunity of expanding this conclusion section to add text to address the issues about the nature of the FireMIP project raised below (the additional text is given in our response to these questions).

4. The scientific methods and assumptions are not clearly outlined, and this make it difficult to judge about their validity. The sections “4. Objective and organization of FireMIP” and “5. Benchmarking and evaluation in FireMIP” are very raw. I would recommend to add a flowchart explaining the conceptual framework of the project objective and organization, and a flowchart explaining the procedure for model benchmarking and evaluation.

Response: Our goal in these sections was to describe the conceptual framework of FireMIP and of the benchmarking that will be performed; we did not want to provide a detailed protocol for the experiments or the benchmarking because these will both be developed during the project itself. We agree that it will be

helpful to provide flowcharts for the experiments and for the benchmarking when we document these protocols. However, we think these are not appropriate in a paper that focuses on giving an overview of fire model development and presenting the state-of-the-art in global fire modelling, as an introduction to the need for a model intercomparison project and for benchmarking current models. We have now tried to make the aim and objective of the current manuscript clearer in the abstract and introduction.

I also think that authors should address the following questions in the text: A) Which of the fire modelling groups are eligible to participate in the project? B) Could any group submit its model for benchmarking and evaluation? C) Were all fire modelling groups invited to participate in the project?

Response: FireMIP is a community initiative rather than a funded project, and has come about through interactions between a large group of fire-modelling groups worldwide. However, participation in this initiative is open to all fire modelling groups, and also to fire scientists who wish to participate in model analysis. One of the purposes of this manuscript is to advertise the FireMIP project to the wider community, to encourage participation. We welcome the chance to make this clear and have added some text in the final section as follows (Lines 494-497):

“FireMIP is a non-funded initiative of the fire-modelling community. Participation in the development of benchmarking data sets and analytical tools, as well as in the running and analysis of the model experiments, is open to all fire scientists. We hope that will maximise exchange of information between modelling groups and facilitate rapid progress in this area of science.”

We have also taken the opportunity to add an invitation to participate in FIREMIP in the acknowledgement section.

D) Is the proposed procedure of model benchmarking and evaluation new and original? E) Are the proposed metrics for model benchmarking and evaluation new and original?

We are sorry that it was not clear that FireMIP will be the first time that benchmarking and evaluation of fire models across standard experiments is carried out. Our benchmarking data is still under development, but as we point out in the manuscript (see also response to reviewer no. 2), the intention is to go well beyond the data sets and metrics that were proposed in Kelley et al. (2013) because we need to have information about other aspects of the fire regime. We have taken the opportunity to make the novelty of the FireMIP initiative clearer by adding a sentence in the last paragraph of the introduction as follows (Lines 139-141):

“There has been no previous attempt to compare fire models across a suite of standardised experiments (model-model comparison) or to systematically evaluate model performance using a wide range of different benchmarks (data-model comparison).”

In addition, we also clarified some of these aspects in revisions to the benchmarking section (e.g., lines 393-402, and 405-412), including a new paragraph (lines 413-418):

“The FireMIP benchmarking system will represent a substantial step forward in model evaluation. Nevertheless there are a number of issues that will need to be addressed as the project develops, specifically how to deal with the existence of multiple data sets for the same variable, how to exploit process understanding in model evaluation, and how to ensure that models which are tuned for modern conditions can respond to large changes in forcing. The answers to these questions remain unclear, but here we provide insights into the nature of the problem and suggest some potential ways forward.”

7. Do the authors give proper credit to related work and clearly indicate their own new/original contribution? I am not sure. It is not clear whether the authors propose new and original procedure and metrics for model benchmarking and evaluation or not.

Response: Please see response to point above.

10. Is the overall presentation well structured and clear? No. The models are reviewed in somewhat chaotic manner. Some models are not mentioned at all.

Response: Fire models have developed in parallel to one another, but there has been some overlap between the approaches taken by different models to representing key processes, which has been the logic behind the current structure in presenting the different models. Indeed (as we show in Tables 1 and 2), some process-descriptions have been adopted by several models –either with minor modification or with tuning because of being coupled to different representations of other aspects of the fire regime. For example, many modelling groups have adopted the human ignition and suppression algorithm, and although the population density thresholds used differ, there is nothing fundamental that distinguishes these treatments. Our goal here was not to describe every single fire model in detail, but rather to outline the major approaches to key processes and in particular to mention models when they introduce fundamentally new approaches (which we now have clarified in lines 197-200). We now have included a mention of the fire module in the IAP RAS CM and citing the Eliseev paper as suggested and have also include reference to two more recent fire model developments.

13. Should any parts of the paper (text, formulae, figures, tables) be clarified, reduced, combined, or eliminated? The sections 4 and 5 need major revision.

We have now tried to explain better the objective of the manuscript in the abstract and introduction. We have also adapted the section 5 and the section conclusions and next steps. See also response to comment about sections 4 and 5 above.

14. Are the number and quality of references appropriate? No. There are no references to the papers of some active fire modelling groups. For example, I did not find references to the papers recently published by Eliseev: Eliseev, A.V., I.I. Mokhov, and A.V. Chernokulsky, 2014: An ensemble approach to simulate CO₂ emissions from natural fires. *Biogeosciences*, 11 (12), 3205-3223, doi 10.5194/bg-11-3205-2014
Eliseev, A.V., I.I. Mokhov, and A.V. Chernokulsky, 2014: Influence of ground and peat fires on CO₂ emissions into the atmosphere. *Doklady Earth Sci.*, v. 459, no. 2, p. 1565-1569, doi 10.1134/S1028334X14120034.

Response: We have now included reference to the fire development in IAP RAS CM as well as two other recent fire developments.

Moreover, there is no one reference on the lines 136-157 where authors review physical controls of fires. This looks strange.

Response: The text describing the physical controls on fire basically summarises what is now “common knowledge” in fire science, and various versions of this description appear in all the major reviews published in the last few years. While we feel it is important to provide this context, there is nothing surprising about this section of text. We could provide multiple references for each statement, but there would be little justification for citing one set of papers over others and any choice (to keep the reference list within reasonable limits) would be arbitrary. For this reason, we prefer not to include references to this section of the text.