Review of "Improving historical trends in the INFERNO fire model using the Human Development Index" by Teixeira et al.
Reviewer: Vincent Verjans

This study incorporates the Human Development Index (HDI) into the INFERNO fire modeling scheme of the JULES land surface model. In this implementation, the HDI decreases linearly the anthropogenic ignitions and the fraction of unsuppressed fires in the model. This modeling initiative is based on the assumption that increased socio-economic development, as approximated by the HDI, leads to more fire-suppressing policies and management. The study evaluates the impact of this new modeling framework using a simulation over the 1997-2016 period at the global scale, where the world is subdivided in 14 regions. Evaluation of the model without the HDI scaling (JULES-INFERNO, JI hereafter) and with the HDI scaling (JULES-INFERNO+HDI, JIH hereafter) is performed with respect to the burned area (BA) product of the Global Fire Emissions Database with small fires (GFED4s).

This study addresses an important and difficult objective: representing in global process-based fire models the influence of socio-economic factors on fire ignition and suppression. The study builds on the notion that societal approaches to fire management vary, and are not only captured by human population density and land use practices. While the socio-economic impact on fire activity depends on a number of complex factors with very large diversity across the world, this study explores if a simple linear scaling using the HDI can improve fire modeling. This is an important step for the fire modeling community, as methods to incorporate socio-economic influence are needed to improve fire modeling. Also, I believe that the modeling procedure of the authors is well-designed, and that their results are a contribution to the field. However, I have a major reservation concerning the presentation of the results. In particular, the presentation is misleading on the benefits of the linear and globally-uniform inclusion of HDI, where conclusions are not always well-supported quantitatively by the results. I note that this concern was already raised in the review of this study in a previous submission to the journal Biogeosciences, and has still not been sufficiently addressed. I detail this Major concern in this review, as well as two Minor comments, and other Specific comments. Nevertheless, I emphasize that, in my view, this study includes sufficient novel work and results that are relevant to the fire modeling community for meriting publication in Earth System Dynamics, but a more transparent presentation is required. Line numbers in this review correspond to the preprint manuscript.

Major comment: Inadequate phrasing and presentation of the manuscript

The entire phrasing of the manuscript should be changed. In its current version, this study tends to be presented as a major breakthrough in capturing socio-economic impacts on fire activity. Results show some improvements for some aspects, but many model deficiencies remain. Some of these deficiencies are even exacerbated in JIH compared to JI. Given this, the study might be better positioned as an exploratory attempt at introducing a simple, globally-uniform linear parameterization, with limited but instructive benefits. A more cautious and quantitatively grounded phrasing would improve the scientific integrity and clarity of the manuscript and help position the work as a useful step toward future refinements. It is important that the phrasing reflects with greater honesty the quantitative evaluation.

In this Major comment, I highlight some of the most obvious examples of this problem in the manuscript. However, I encourage the authors to revise fundamentally the writing of the manuscript. These examples are presented in the order they appear in the manuscript, and are not in order of

importance.
Title: The notion of "improving" should be removed.
l14: "way to improve fire model performance": this statement is too general and vague. Since there is no general improvement of performance, the aspects that are improved need to specified.
l16: A sentence in the abstract should be added to explain that a linear and globally-uniform inclusion of HDI as a simple approximation for socio-economic factors is a step forward but insufficient in many aspects.
l211: "including socio-economic factors": here and everywhere in the manuscript, the authors should not write that they include socio-economic factors, but that they include the HDI. The former wording suggests a more complex implementation than what is truely done, i.e., only including HDI. Nevertheless, this wording is repeatedly used (e.g., l278, l285, l389, l458, l465, l514, and many more).
l254: "This evidence highlights that HDI can be used as an indicator of the role socio-economic factors play in mitigating fire activity". This sentence is too strong. This evidence only shows that HDI can regionally capture part of the variability in BA.
l264: "aligning the model more closely with observations" should be: aligning the model more closely with observations in terms of global mean dependence on HDI.
l259: "leading to improvements over North America, Europe and Asia, as shown in Figure 8". Figure 8 only shows the bias, so this statement should be: leading to reduced bias (...). In general, it is important to be more precise in the wording, rather than using general terms not supported by the Figures and/or numbers that are referred to.
l293: Here, the text only mentions the part of the histogram where JIH performs better than JI (i.e., BA fractions between 0.7 and 1.0). The authors avoid addressing the BA fractions <0.5, where JI strongly outperforms JIH, and which represent the large majority of the fire occurrences.
l305: "In EURO, the inclusion of socio-economic factors better represents both small and moderate burnt area fractions". This is not true. Fig. 9g does not show a better performance of JIH compared to JI.
l325: "Nevertheless, discrepancies remain in some regions" should be changed to: Nevertheless, discrepancies are exacerbated in some regions.
l326: "the inclusion of HDI represents a significant advancement": based on which metric is this strong statement made?
l345: Here, the authors list regions of larger bias of JIH, but they omit that the bias is also larger for the global scale.
l348: "a reduction of bias and RMSE in regions where improvements are most needed": based on which criteria do the authors estimate that some regions are more in need of improvements than others?
l354: "improving the representation of global burnt area variability": this statement is false, as is clearly shown in Fig. 10a. The reason for this false statement is that the metric used by the authors (STD/STD_{GFED4s}) is inadequate. That is because the standard deviation is influenced by the magnitude of the trend. To provide an adequate measure of inter-annual variability, the authors should compute the standard deviation after removal of a linear trend. In this case, I am almost certain (based on a visual analysis of Fig. 10a) that inter-annual variability at the global scale is larger in JI than in JIH.
l396: This paragraph illustrates that there are also compensating biases in JIH that lead to a deceptively better skill of JIH than JI for some aspects such as the global trend. However, while the authors often use the wording "compensating biases" when describing the performance of JI, they never use this wording to describe the JIH performance.
l485: "an improved representation of the relationship between burnt area and HDI" should be: an

improved representation of the globally-averaged linear relationship between burnt area and HDI.
l489: "The observed linear relationship between burnt area and HDI". There is no observed linear relationship between HDI and BA (Fig. 3). It is approximated as linear by the authors. The fact that the posterior distribution for the slope is significantly negative does not mean that the relationship is linear. It only means that if we assume a linear relationship, then there is a significant non-zero slope.
l499: The authors omit to write explicitly that JIH has an increased bias compared to JI at the global scale (same as comment above about l345).
l511: "discrepancies against observations remain in JULES-INFERNO+HDI" should be: discrepancies against observations are exacerbated in JULES-INFERNO+HDI.
l513: "where the model continues to underpredict medium and large fire sizes" should be: where the model underpredicts medium and large fire sizes more strongly.
l517: "misrepresents the observed positive burnt area trends found in TENA, MIDE and SEAS": the region BONA should also be listed here.
l583: "Although this could be seen as a negative impact" should be: Although this further exacerbates the under-estimation of inter-annual variability in JULES-INFERNO+HDI (...).
l621: "provides a simple and linear representation of these effects": replace representation by approximation.
l621: "This leads to an improvement in model performance, especially in developed regions." This statement is debatable. Again, there is no general increase in performance. So, such a statement should be more specific on which aspects are improved.

Minor comment 1: Some methodological aspects are unclear

(a) Figure 2. It is not clear how the FWI is linearly regressed out of the BA. Is this regression performed at the level of individual grid cells or regions or globally? Does it use every monthly BA value at each grid cell, or is it based on the climatology of the FWI and BA? Please provide more detail, and also a figure of the linear regression in the Appendix, along with an R^2 statistic.

(b) The Bayesian fitting procedure of Figures 3 and 6 should be better explained.
l126: "δBA having a log-normal distribution with mean of zero and standard deviation of ten". If δBA has a log-normal distribution, this means that $\log(\delta BA)$ follows a Normal distribution, and thus that δBA is constrained to be positive. But by comparing the legend in Fig. 3 with Eq. (1), it appears that the posterior mean of δBA is -6.57. There is an inconsistency in the explanation of the priors and/or Eq. (1), which needs to be corrected.
Figures 3 and 6: since the quantity of interest is δBA, please show the posterior distributions of this coefficient in the main manuscript instead of the Appendix.
Figures 3 and 6: Please specify if the grey points of the scatter plot represent all monthly BA values at all grid cells of GFED4s (Fig. 3) and of the models (Fig. 6). I believe so, but it is not explained.
l128: Why did the authors chose to represent the posterior uncertainty with 145 posterior samples? This choice seems arbitrary.
l123: "optimization over a normal posterior distribution": do the authors mean a normal likelihood?

(c) The analysis of Figure 9, is very qualitative. I recommend that, in each subplot of Figure 9, the authors provide a quantiative metric of the fit of the JI and JIH histograms to the

GFED4s histogram, for example the Wasserstein distance. Their analysis (from l283 to l331) would benefit from a more quantitative description of the performance.

(d) Too many trends are analyzed in Figure 13 (11 model configurations times 15 spatial entities = 165 trend values). I recommend to show in Figure 13 only the trend values that are significantly different from zero, and to limit the analysis (l434 to l478) to the significant trends only.

Minor comment 2: Referencing of literature
The referencing in the manuscript does not currently meet the standards expected for Earth System Dynamics. Several statements lack appropriate citations where references are clearly needed, while others cite sources that do not adequately support the claims being made. In some cases, multiple references (often more than four) are grouped at the end of a paragraph to justify the entire content, which makes the specific contributions of individual studies unclear. It would be more informative to cite specific examples from the literature in direct connection with the relevant claims. Additionally, some important contributions from the existing literature that are highly relevant to this study are missing entirely. In this Minor comment, I try to provide specific instances where the referencing could be improved, and I hope that this will help the authors to better reference existing literature in the manuscript.

l21: "decline of 1.27% per year": requires a citation.
l24: "Climate is a key factor that also influences fire activity (Archibald et al., 2010; Andela et al., 2017; Jones et al., 2022; Kelley et al., 2019)." I do not think that 4 citations are needed to state this well-know fact. If these 4 studies are relevant here, then please indicate the specific aspects of these studies that are important to highlight.
l30: I believe that it is critical to cite the work of Marlon et al. (2008) here. In particular this study shows how anthropogenic factors have affected changes in fire activity over multi-decadal to multi-centennial time scales.
l36: I believe that, in this paragraph, it is critical to cite the work of Forkel et al. (2019). Their quantification in a data-driven framework of the anthropogenic versus climate drivers of fire activity is very relevant to this study. (I recommend the authors to have a look at Fig. S13a,b, which they might find interesting).
l42: "However, most CMIP6 models do not adequately account for these suppression mechanisms, resulting in an overestimation of burned area and fire-related carbon emissions." This statement is false. Please see Figure 2 of Li et al. (2024). It is clear that most CMIP6 models represent well global total BA and fire C emissions. And they tend to under-estimate rather than over-estimate these quantities.
l48: At this point of the introduction, I believe that it would be valuable to shortly describe how existing fire models quntify changes in fire activity from climate change drivers verus anthropogenic influence. With regards to this aspect, referencing the study of Burton et al. (2024) would be relevant.
l51: At this point of the introduction, I believe that it would be valuable to shortly explain that simulating fire accurately is also important for climate projections, because of two-way feedback processes between climate and fire. With regards to this aspect, referencing the study of Verjans et al. (2025) would be relevant.
l55: "The HDI has been used in various studies to better understand the socio-economic impacts on the Earth System (ES) (Türe, 2013; Hickel, 2020; Roy et al., 2023)." If these 3 studies are relevant here, then please indicate the specific aspects of these studies that are important to highlight.

l57-61: I find this paragraph confusing because the notions of inter-annual variability and fire activity are used interchangeably. Please note that Chuvieco et al. (2021) only focus on inter-annual variability, and re-phrase the paragraph accordingly.
l63: “However, their approach was limited to agricultural fires and did not account for broader human factors in fire management.” This statement is false. Li et al. (2013) use a GDP-based parameterization that is not limited to agricultural fires. Please revise this paragraph.
l63: I believe that it is critical to cite the work of Perkins et al. (2024) here. In particular, please discuss recent more sophisticated attempts to incorporate anthropogenic drivers in fire modeling.
l77: “Several studies have shown that in developed regions, land and fire management policies play a more significant role in controlling fire ignitions than other human behaviours (Nikolakis and Roberts, 2022; Jacobson et al., 2022; Ford et al., 2021; Curt and Frejaville, 2018; Carreiras et al., 2014; Mourão and Martinho, 2014).” If these 6 studies are relevant here, then please indicate the specific aspects of these studies that are important to highlight.
l491: “For instance, the gross national income index indicates that higher HDI regions typically have more funding available for fire prevention and suppression efforts (Rideout et al., 2017).” This specific fact is not mentioned in Rideout et al. (2017).
l493: “Similarly, the life expectancy index suggests that these governments are more likely to implement policies aimed at mitigating the negative impacts of fire on their population (Rizzo and Rizzo, 2024).” This specific fact is not mentioned in Rizzo and Rizzo (2024).
l494: “ Additionally, the education index highlights that educational initiatives can enhance community awareness and preparedness regarding fire risks and environmental stewardship (Prestemon et al., 2010).” Please be careful here, as this statement suggests a very general fact, while the study of Prestemon et al. (2010) focuses only on the state of Florida.
l540: “Although HDI does not encompass explicitly the impacts of fire management policies, these results are consistent with other studies, which show that for developed regions, land and fire management policies have a greater role than other human behaviours in controlling ignitions (Nikolakis and Roberts, 2022; Ford et al., 2021; Jacobson et al., 2022; Carreiras et al., 2014; Mourão and Martinho, 2014).” If these 5 studies are relevant here, then please indicate the specific aspects of these studies that are important to highlight.
l552: “Several authors have also shown that declines in burnt area in the Mediterranean have occurred irrespective of increases in fire weather, as well as extensions to the fire weather season length, which is attributed to increased fire prevention and in combating and mitigating fire impacts (Jones et al., 2022; Urbieta et al., 2019; Carreiras et al., 2014; Mourão and Martinho, 2014).” If these 4 studies are relevant here, then please indicate the specific aspects of these studies that are important to highlight.
l560: The facts about fire in the Amazonia region stated here should use citations at the end of the specific sentences, rather than grouping 3 citations together at the end of the paragraph.
l573: “However, INFERNO has been developed for Earth System Modelling resolutions and timescales, and it is not expected to be able to capture the representation of the processes that drive large and severe fires”. Please be careful here, because this wording suggests that this is a limitation inherent to all Earth System Models. However, this limitation is rather due to the use of $\overline{BA_{PFT}}$ in Eq. (2). Please cite counter-examples, for example Lasslop et al. (2014).
l595: “In addition, biases in the underlying vegetation can significantly impact modelled burnt area”. Please also refer to the work of Forkel et al. (2019) here, as they demonstrated some widespread shortcomings of fire models in capturing the sensitivity of BA to leaf area index and plant productivity.
l603: “it is known that JULES vegetation has few needle-leaf trees across the boreal regions compared to observations”. Please provide a citation here.

Specific comments

l1: "Earth System Models (ESM), have struggled to reproduce the historical decline in burnt area": this statement is too crude. See for example Fig. 1 of Teckentrup et al. (2019), and it is also shown by Li et al. (2024) that most CMIP6 models capture the 1850-2010 trend.
l4: "formulation" should be plural.
l5: Specify period of the trend.
l7: Change "reflects" to aims to reflect.
l7: Change "and, in turn" to , which in turn.
l9: Specify: reduces biases in annual burnt area for some regions, particularly (...)
l15: Change "human-environment" to human-fire.
l18: Specify: climate change and variability.
l60: This is the first time that fire is mentioned as a land management tool. Up to here, the Introduction only focuses on fire suppression. This notion of land management needs to be introduced before.
l68: The objective is rather to model the influence of human populations on fire activity.
l97: Comma is missing after "Section 4".
l101: Here and in the remainder of the manuscript, correlation values should not be given in %.
Figure 2: When testing for significance, did the authors apply a correction for false discovery rate? If not, this is needed here (please see Wilks, 2016).
Figure 2 legend: "significant with a 95% confidence level" should be: significant at the 5% level.
l107: In the analysis of Fig. 2b, the first and most important aspect to focus on is that over most of the globe, the correlation between deweathered BA and HDI is not significant. The analysis should then only focus on the areas with statistical significance.
l107: "Figure 2 shows the spatial correlation coefficient": I believe that Figure 2 shows the temporal correlation coefficient.
l114: Specify: a strong positive correlation with FWI.
l130: Typo: this method shows with an s.
l130: Specify: the observations show a log-linear decline.
l130: Specify: with a posterior mean slope of -6.57.
l140: If this is correct, specify: $\overline{BA_{PFT}}$ is the average burnt area per fire for each PFT.
l142: "This decouples the fire spread stage from local meteorology": I believe that this is not entirely true, because the F_{PFT} depends on local meteorology.
l145: Please remove "significantly".
l186: Typo: dataset
Figure 4: This figure should show a second sub-panel of f_{NS} as a function of PD at different HDI levels.
l202: The LIS-OTD climatology provides total lightning flash density. What parameterization is used to convert total flash density to cloud-to-ground flash density? Please specify this important aspect in the text.
l211: "(...) for the ignitions and suppression of fires. This is reflected in the $\overline{BA_{PFT}}$ values (...)" should be rephrased to (...) for the ignitions and suppression of fires, which affects the $\overline{BA_{PFT}}$ values (...).
Table 1: Are these all the PFTs of the JULES model? Please specify in the caption.
Section 2.4: In my view, this section can be moved to the Appendix.
Equations (9),(10,(11): The small n should be capital N.
l245: "additional noise" is inappropriate wording here, and should be replaced by: residual variability.

l248: "-6.57 ($\%^{-1}$)": I believe that there is an error in the units here.
l249-254: This paragraph should focus only on those regions where the correlation between HDI and deweathered BA is significant.
l260: The mention of JULES-INFERNO+HDI in this sentence is confusing. First compare with observations, then compare with JIH. Please restructure this paragraph.
l270: "The 2-D cross-correlation was used to determine what is referred to as spatial correlation between the model experiments and the observation data." This is commonly called: pattern correlation.
l277: There are two many closing parentheses here, and "1 e" should be specified as Fig. 1e.
l279: Remove "high prosperity", and just use "high values of HDI".
l279: "over North America, Europe and Asia". These regions are too broad. Please be more specific (e.g., Iberian peninsula, etc.).
Figure 8: Please show a third sub-panel that shows the difference between the absolute bias of JIH and the absolute bias of JI.
l283: In the entire analysis of Fig. 9, I find it confusing that the authors always discuss "fire sizes", whereas Fig. 9 shows the BA fraction. In any given month, the BA fraction is the product of the number of fires times the mean BA per fire, then normalized by the total grid cell area. So, how do the authors know that changes in BA fraction are one-to-one related to the fire size? Why could this not reflect a larger number of fires? Please explain why Fig. 9 can be interpreted as fire size distributions.
Figure 9: As mentioned in my Minor comment (c), please provide a quantitative metric (e.g., Wasserstein distance) and use this metric in the comparison of JI versus JIH performance with respect to GFED4s.
Also, please use the same span for the y-axis across all the sub-panels. This would facilitate the analysis of Fig. 9 for the reader.
l299: " the frequency of burnt areas is noticeably reduced". As far as I understand, the sum of the frequencies should always be 100%. Thus, should this be: the frequency of non-zero burnt areas?
l308: "In regions such as Central America (CEAM) and Southern Hemisphere South America (SHSA), GFED4s shows broader frequency distributions (...)". This is not clear, particularly for CEAM. By using a common y-axis span for all sub-panels (see comment above), this could appear more clearly.
l317: "East Asia Asia" should be Southeast Asia.
l317: This sentence states that Australia and New-Zealand is an Asian region, which is not true.
l335: Specify: temporal Pearson correlation.
l336: Specify: a simple log-transformed linear regression with time as predictor.
l350: Specify: to impact negatively regions which experience high levels of burning inter-annual variability.
l354: Throughout the study, the metric STD/STD_{GFED4s} is misinterpreted (see Major comment).
l355: I recommend not using the metric $RMSE_{UB}$ since I do not see the point of analyzing unbiased model error. The resulting information is already conveyed by temporal correlation. If there is a good reason to analyze this metric, it should be better explained in the text.
l359: "The improvements from JULES-INFERNO+HDI in regions such as TENA, NHAF, and SHAF have a greater impact on the global metrics than the reduced performance seen for regions such as CEAM, NHSA, SHSA, EURO, and MIDE." I do not understand why the authors make this argument, nor how it should be interpreted. Please explain this better or remove.
Figure 11: I find this figure very informative. However, to focus on the most important aspects, I recommend to show only the three most important metrics: Relative bias, RMSE, and Correlation. The other metrics can be shown in the Appendix. Reducing the number of metrics would also allow

to show results for all the regions. Additionally, please use a log-scale for RMSE.
l370: It would be insightful to show this quantitative dependence of impact from variations in population at different HDI levels in a Figure.
Figure 12: Please provide the pattern correlation between JI and GFED4s in (b) and of JIH and GFED4s in (c). Also, please add a second row to the figure, showing the same field but with colors saturated at lower levels (e.g., -0.5 to 0.5) so that trends of lower magnitude do not simply appear as grey.
l383: I find the wording "the antecedent precipitation-burned area response" somewhat confusing. Could this be changed to: the antecedent precipitation-driven vegetation build-up? Or something similar?
l385: "it should not be expected for the model to perform well in regions where this precipitation-burned area coupling can be dominant". It has now been extensively demonstrated that precipitation is the main driver of burned area. If INFERNO is not able to reproduce this dependency, then this would be a major problem. Do the authors refer here to the vegetion-build up caused by precipitation? If so, this should be better explained. As it is, this sentence could be interpreted as an unacceptable model limitation.
l394-395: This sentence repeats the exact same information as the previous sentence.
l403: "While JULES-INFERNO+HDI tends to enforce decreasing trends, this only happens in four regions out of 14 (i.e., TENA, SHAF, MIDE, and SEAS)." This wording is highly confusing. Please simply this sentence. For example, simply: JULES-INFERNO+HDI shows decreasing trends in all 14 regions, including those 4 with positive trends in GFED4s (i.e., TENA, SHAF, MIDE, and SEAS).
l410: Here, and in other sentences where the inability of INFERNO to simulate large fires is mentioned, the authors should provide a better explanation. They should explicitly refer to the use of $\overline{BA_{PFT}}$ in Eq. (2). They should also explain why I_T and F_{PFT} in Eq. (2) cannot compensate for the use of $\overline{BA_{PFT}}$ to generate large values of BA_{PFT}. The reasons for this inability to simulate large fires are still unclear to me.
l431: If I understand correctly, "the underlying land surface state" should be replaced by: the 1990 land surface state.
l433: Typo, this should be: are given in Supplementary Figure A3.
Figure 13: As explained in my Minor comment (d), please only show the symbols corresponding to significant trends. This would improve the clarity of the figure. This would also make the analysis of the Figure more concise and to-the-point. Also, it would be insightful to add a third row showing the difference between trends in JIH and JI (i.e., sub-panel b minus sub-panel a).
l472: "the role of anthropogenic drivers (land use and population density)". Here and elsewhere in the manuscript, I find this naming confusing because HDI is also an anthropogenic driver.
l508: Please change "resulted" to the present tense to preserve consistency of tenses in the manuscript.
l520: "While JULES-INFERNO+HDI tends to strengthen decreasing trends, this only happens in four regions out of 14 (TENA, SHAF, MIDE, and SEAS)". This sentence is not true, as JIH has a more strongly decreasing trend than GFED4s in more than these 4 regions. Please rephrase.
l524: "For example, INFERNO was not designed to capture the dynamics of large, severe fires that dominate fire regimes in some regions". Same comment here as for l410. Please explain this limitation of INFERNO better.
l530: "The results of these experiments show that including socio-economic impacts on fire results in the burnt areas trends being dominated by socio-economic drivers through a reduction in the contribution from climate drivers, especially from temperature and precipitation." This sensitivity analysis is derived under the hypothesis that the (1-HDI) factor in Eqs. (3), (4) is the correct representation. Another model form would lead to another sensitivity to climate drivers. As such,

I believe that the important conclusion here is rather that accounting for HDI (or other socio-economic factors) alters the model sensitivity to climate drivers. This would change the sensitivity of modeled fires to climate change.

l535: "For example, the inclusion of socio- economic factors reduces the role of temperature in driving trends (e.g., increase for TENA, EURO, CEAS, and AUS), as well as by changing the behaviour that climate drivers have in burnt area trends (e.g., MIDE, NHAF, and SEAS)." I do not understand the difference between the 1st and 2nd parts of the sentence. Could they not be merged in a single statement about changes in the sensitivity to climate drivers?

l557: "resulting in a decrease in burnt area". Should this not be: resulting in a weaker impact on burnt area?

l557: Note that it is hard to see any difference for BONA, CEAM, and NHSA in Fig. 13 between JI and JIH.

l565: The word "also" is repeated twice in this sentence.

l566: Typo, "producing" should be produce.

l581: "While this improves INFERNO performance over regions such as TENA and CEAM". Please be more specific about these improvements.

l586: "Although socio-economic factors are included in JULES-INFERNO+HDI, the HDI dataset provides information mainly at a national level." Why does this sentence start with Although? I do not see the link between the 1st and the 2nd part of the sentence.

l604: In this sentence, "these regions" refer to boreal regions that are mentioned in the previous sentence. And then, in the next sentence, examples are given for India and southwest Russia. Please re-phrase these sentences in order to be more consistent in the argumentation.

l611: "continuity imposed by changes in socioeconomic drivers": I do not understand this. Please clarify.

l624: Here is the last case where the notion of including "socio-economic factors" into INFERNO should be replaced by just HDI. I just wanted to remind the authors to do this replacement everywhere needed in the manuscript (see Major comment).

l642: It would be beneficial to add a sentence in the conclusion that explicitly insists on the fact that more work and, most likely, more complex parameterizations are needed to capture the complexity of socio-economic impacts on fire activity in ESMs.

l653: Typo here: ranging.

l662: Please clarify what is meant by "normalised" in this sentence.

l664: "spatial resolution of 5 arc-minutes". So, please explain why the resolution shown in Figure 1 appears much coarser.

l686: "atmospheric forces" should be: atmospheric forcing.

References

Chantelle Burton, Seppe Lampe, Douglas I. Kelley, Wim Thiery, Stijn Hantson, Nikos Christidis, Lukas Gudmundsson, Matthew Forrest, Eleanor Burke, Jinfeng Chang, Huilin Huang, Akihiko Ito, Sian Kou-Giesbrecht, Gitta Lasslop, Wei Li, Lars Nieradzik, Fang Li, Yang Chen, James Randerson, Christopher P. O. Reyer, and Matthias Mengel. Global burned area increasingly explained by climate change. *Nature Climate Change*, 14(11):1186–1192, October 2024. ISSN 1758-6798. doi: 10.1038/s41558-024-02140-w. URL `http://dx.doi.org/10.1038/s41558-024-02140-w`.

Emilio Chuvieco, M. Lucrecia Pettinari, Nikos Koutsias, Matthias Forkel, Stijn Hantson, and Marco Turco. Human and climate drivers of global biomass burning variability. *Science of The Total*

Environment, 779:146361, July 2021. ISSN 0048-9697. doi: 10.1016/j.scitotenv.2021.146361. URL `http://dx.doi.org/10.1016/j.scitotenv.2021.146361`.

Matthias Forkel, Niels Andela, Sandy P. Harrison, Gitta Lasslop, Margreet van Marle, Emilio Chuvieco, Wouter Dorigo, Matthew Forrest, Stijn Hantson, Angelika Heil, Fang Li, Joe Melton, Stephen Sitch, Chao Yue, and Almut Arneth. Emergent relationships with respect to burned area in global satellite observations and fire-enabled vegetation models. *Biogeosciences*, 16(1):57–76, January 2019. ISSN 1726-4189. doi: 10.5194/bg-16-57-2019. URL `http://dx.doi.org/10.5194/bg-16-57-2019`.

Gitta Lasslop, Kirsten Thonicke, and Silvia Kloster. Spitfire within the mpi ¡scp¿e¡/scp¿arth system model: ¡scp¿m¡/scp¿odel development and evaluation. *Journal of Advances in Modeling Earth Systems*, 6(3):740–755, August 2014. ISSN 1942-2466. doi: 10.1002/2013ms000284. URL `http://dx.doi.org/10.1002/2013MS000284`.

F. Li, S. Levis, and D. S. Ward. Quantifying the role of fire in the earth system – part 1: Improved global fire modeling in the community earth system model (cesm1). *Biogeosciences*, 10(4):2293–2314, April 2013. ISSN 1726-4189. doi: 10.5194/bg-10-2293-2013. URL `http://dx.doi.org/10.5194/bg-10-2293-2013`.

Fang Li, Xiang Song, Sandy P. Harrison, Jennifer R. Marlon, Zhongda Lin, L. Ruby Leung, Jörg Schwinger, Virginie Marécal, Shiyu Wang, Daniel S. Ward, Xiao Dong, Hanna Lee, Lars Nieradzik, Sam S. Rabin, and Roland Séférian. Evaluation of global fire simulations in cmip6 earth system models. *Geoscientific Model Development*, 17(23):8751–8771, December 2024. ISSN 1991-9603. doi: 10.5194/gmd-17-8751-2024. URL `http://dx.doi.org/10.5194/gmd-17-8751-2024`.

J. R. Marlon, P. J. Bartlein, C. Carcaillet, D. G. Gavin, S. P. Harrison, P. E. Higuera, F. Joos, M. J. Power, and I. C. Prentice. Climate and human influences on global biomass burning over the past two millennia. *Nature Geoscience*, 1(10):697–702, September 2008. ISSN 1752-0908. doi: 10.1038/ngeo313. URL `http://dx.doi.org/10.1038/ngeo313`.

Oliver Perkins, Matthew Kasoar, Apostolos Voulgarakis, Cathy Smith, Jay Mistry, and James D. A. Millington. A global behavioural model of human fire use and management: Wham! v1.0. *Geoscientific Model Development*, 17(9):3993–4016, May 2024. ISSN 1991-9603. doi: 10.5194/gmd-17-3993-2024. URL `http://dx.doi.org/10.5194/gmd-17-3993-2024`.

Jeffrey P Prestemon, David T Butry, Karen L Abt, and Ronda Sutphen. Net benefits of wildfire prevention education efforts. *Forest Science*, 56(2):181–192, 2010.

Douglas B Rideout, Y Wei, A Kirsch, and N Kernohan. Starfire: Strategic budgeting and planning for wildland fire management. *Park Science*, 32(3):34–41, 2017.

Luciana V Rizzo and Maria Cândida FV Rizzo. Wildfire smoke and health impacts: a narrative review. *Jornal de Pediatria*, 101:S56–S64, 2024.

Lina Teckentrup, Sandy P Harrison, Stijn Hantson, Angelika Heil, Joe R Melton, Matthew Forrest, Fang Li, Chao Yue, Almut Arneth, Thomas Hickler, et al. Response of simulated burned area to historical changes in environmental and anthropogenic factors: a comparison of seven fire models. *Biogeosciences*, 16(19):3883–3910, 2019.

Vincent Verjans, Christian L. E. Franzke, Sun-Seon Lee, In-Won Kim, Simone Tilmes, David M. Lawrence, Francis Vitt, and Fang Li. Quantifying co_2 forcing effects on lightning, wildfires, and climate interactions. *Science Advances*, 11(7), February 2025. ISSN 2375-2548. doi: 10.1126/sciadv.adt5088. URL `http://dx.doi.org/10.1126/sciadv.adt5088`.

DanielS Wilks. “the stippling shows statistically significant grid points”: How research results are routinely overstated and overinterpreted, and what to do about it. *Bulletin of the American Meteorological Society*, 97(12):2263–2273, 2016.