

Current model capabilities and future predictive pathways of wildfires under a changing climate

Corresponding Author: Dr Cynthia H. Whaley

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

I read this as a generalist (rather than modelling expert) and found the paper extremely useful and thoroughly researched. It is well written and fills an important niche in the global wildfire literature.

I have not major criticisms and only a few suggestions.

1. The authors talk a lot about various earth system components that interact in various ways, and they have a nice diagram of some of these factors. I think an additional diagram showing linkages would help appreciate interactions and feedback loops, especially if the linkages were colour coded to show relationship that are routinely or rarely considered in LSM and ESM. I appreciate these relationships are indicated in the Figure 1, but reading the text there are numerous climate, chemical and biological variables that are discussed that could be put into an influence diagram to highlight the complexity of global wildfire modelling.

2. The authors talk a lot about model predictions and uncertainties, but can they state clear take homes in a table or a text box. What aspects of fire regime on Earth are most likely and least likely to change. It would be impactful to have this information available at a glance.

3. The focus of the review is modelling future global fire activity, but can the authors indicate how the models are reinforcing (or counteracting) current empirically derived trends in wildfire activity. If the model predictions are reinforcing current trends this reciprocally strengthens both the inference of current trend analyses and faith in model predictions.

Reviewer #2

(Remarks to the Author)

Summary:

This manuscript presents a comprehensive review of current modelling approaches used to project wildfire activity under climate change, including process-based land surface and Earth system models (LSMs/ESMs), as well as statistical and machine learning (ML) approaches. The authors synthesize recent literature to compare model structures, key drivers, and projected changes across regions and climate scenarios, and highlight major uncertainties and emerging directions, including hybrid modelling frameworks and the integration of AI techniques.

The topic is timely and highly relevant. The manuscript is clearly written, well-structured, and provides a very useful and comprehensive synthesis of a rapidly evolving field.

In my opinion, it will be of clear interest to researchers working on wildfire-climate interactions and Earth system modelling.

General assessment:

Overall, this is a strong and thorough review that successfully compiles a large and diverse body of literature. The manuscript is particularly valuable in its impressive breadth of coverage, including extensive supporting material, and in its effort to bridge process-based and data-driven modelling paradigms. However, in its current form, the manuscript remains largely descriptive. While it provides an extensive overview of existing approaches and results, it falls somewhat short of delivering the level of critical synthesis and conceptual advancement that would make it a clear "significant advance" for Communications Earth & Environment. Strengthening the analytical and comparative aspects of the review would substantially enhance its impact. In this sense, the availability of detailed model inventories (e.g. in the Supplementary

Information) suggests that a more explicit comparative synthesis across modelling frameworks is both possible and desirable.

From the perspective of impacts and decision-relevant assessments, the manuscript would also benefit from clearer guidance on how to interpret and handle these uncertainties. In particular, a more explicit distinction between critical vs. secondary sources of uncertainty, and their implications for projection credibility and usability, would be extremely valuable for both researchers and end users.

Major comments:

1. Limited critical synthesis across modelling frameworks

A central weakness of the manuscript for Communications Earth & Environment is that it primarily summarizes the literature rather than performing a more critical evaluation. The discussion of LSMs, ESMs, and ML approaches is thorough and informative but largely parallel. In this sense, I am missing a more critical perspective, including a more explicit and direct comparison of their relative strengths, weaknesses, and domains of applicability.

The manuscript does not clearly articulate under which conditions one modelling framework provides more reliable or informative projections than others.

In this sense, I encourage the authors to include a more explicit cross-framework synthesis, identifying:

- * where models agree vs. disagree
- * which sources of uncertainty are most consequential
- * which approaches currently appear most promising for specific applications (e.g. long-term projections vs. near-term forecasting...)

2. Example of critical analysis: Physical consistency and implications for compound events

Following the last point, an important aspect that deserves more explicit discussion is the difference in physical consistency between process-based and machine learning approaches.

Process-based models (LSMs/ESMs), by construction, enforce internally consistent relationships among key variables (e.g. energy, moisture, vegetation dynamics), ensuring that simulated states remain physically plausible under both present and future conditions. In contrast, statistical and machine learning models, while often highly skillful in reproducing historical patterns, do not guarantee such internal consistency unless explicitly constrained. This distinction becomes particularly important in the context of compound and extreme fire-conducive events, where the joint behaviour of multiple variables (e.g. temperature, humidity, wind, fuel moisture) is critical, particularly under non-stationary conditions and in extrapolation regimes relevant for future climate projections. Models lacking inherent physical constraints may capture correlations in the training data but produce unrealistic or physically inconsistent combinations under novel climate conditions.

While the manuscript mentions limitations of ML approaches in terms of extrapolation and missing feedbacks, a more explicit discussion of physical consistency and its implications for compound events and extreme regimes would strengthen the comparative assessment of modelling frameworks and provide valuable guidance for future developments, particularly in the context of hybrid and physics-informed approaches.

Furthermore, this also reinforces the potential importance of physics-informed machine learning approaches, which may help bridge the gap between predictive skill and physical consistency.

3. Lack of quantitative or structured comparison

Closely related to the previous point, the manuscript would benefit from a more structured and, where possible, quantitative comparison of modelling approaches.

Examples of possible improvements:

- * summary tables comparing key features of LSMs, ESMs, and ML frameworks (e.g. feedback representation, spatial scale, computational cost, ability to extrapolate)
- * more systematic comparison of projected changes (ranges, sensitivities) across studies

At present, many statements remain qualitative, making it difficult for the reader to assess the relative robustness of different approaches. In my opinion, the Supplementary Information already compiles detailed model descriptions and classifications; distilling this information into comparative tables or synthesis figures in the main text would substantially improve accessibility and impact.

4. Prioritization of uncertainties, which ones are critical and which ones secondary?

The manuscript identifies a wide range of uncertainties (climate forcing, vegetation dynamics, lightning, human activity, model structure), but these are presented largely as a list.

What I am missing is a stronger effort to prioritize uncertainty sources and to clearly distinguish between dominant vs. secondary drivers and structural vs. "parametric" uncertainties, including an assessment of which uncertainties most strongly limit predictive skill and the credibility of the projected changes. In my opinion, providing such prioritization would significantly improve the manuscript's usefulness and impact.

5. Limited conceptual advance

While the review is comprehensive, it does not clearly introduce a new conceptual framework or synthesis that changes how the reader understands the field. For a journal of this level, it would be desirable to include:

- * a more explicit conceptual synthesis (e.g. a unifying framework linking climate, fuels, and human processes across modelling approaches)
- * or a clearer articulation of key paradigms or regimes governing wildfire projections

At present, the manuscript reads more as an extensive and well-executed compilation than as a “transformative” synthesis.

6. Treatment of human-fire interactions

The role of human activity is acknowledged throughout the manuscript, but it is often treated as one of several uncertainties rather than a central driver. Given the strong evidence that land use, suppression, and ignition patterns dominate fire regimes in many regions, the manuscript would benefit from a more integrated treatment of human-fire interactions across all modelling frameworks and a clearer discussion of how different models represent (or fail to represent) these processes in a realistic way.

7. Towards actionable projections: Link between model evaluation and projections

The manuscript highlights the need for actionable projections, but the discussion remains relatively general. It would be valuable to more explicitly link model capabilities and uncertainties to decision-relevant metrics and discuss how current limitations affect the usability of projections for fire management and adaptation planning. In addition, the manuscript compiles a substantial body of literature on both model evaluation and future projections (as also reflected in the Supplementary Information). However, the connection between model skill in reproducing historical fire behaviour and confidence in future projections remains underdeveloped. A more explicit discussion linking model evaluation, identified biases, and projection credibility would significantly strengthen the synthesis and help readers assess which projections are most trustworthy.

Minor comments

Some sections (particularly in the modelling framework overview) could potentially be slightly condensed to improve focus on synthesis.

Figures are helpful; however, additional schematic or summary figures (e.g. comparing frameworks or uncertainty sources) could further enhance clarity. However, this is in my view totally secondary.

Open Access This Peer Review File is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

In cases where reviewers are anonymous, credit should be given to 'Anonymous Referee' and the source.

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons

license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Reviewer comments in *blue italic*

Author response begins with [AR]

[AR]: A quick note that the original manuscript was ~6700 words with 6 display items, and while we added to the manuscript to address the reviewer comments below, we also edited further for conciseness. The revised manuscript is ~6200 words with 8 display items (7 figures and 1 text box).

Reviewers' comments:

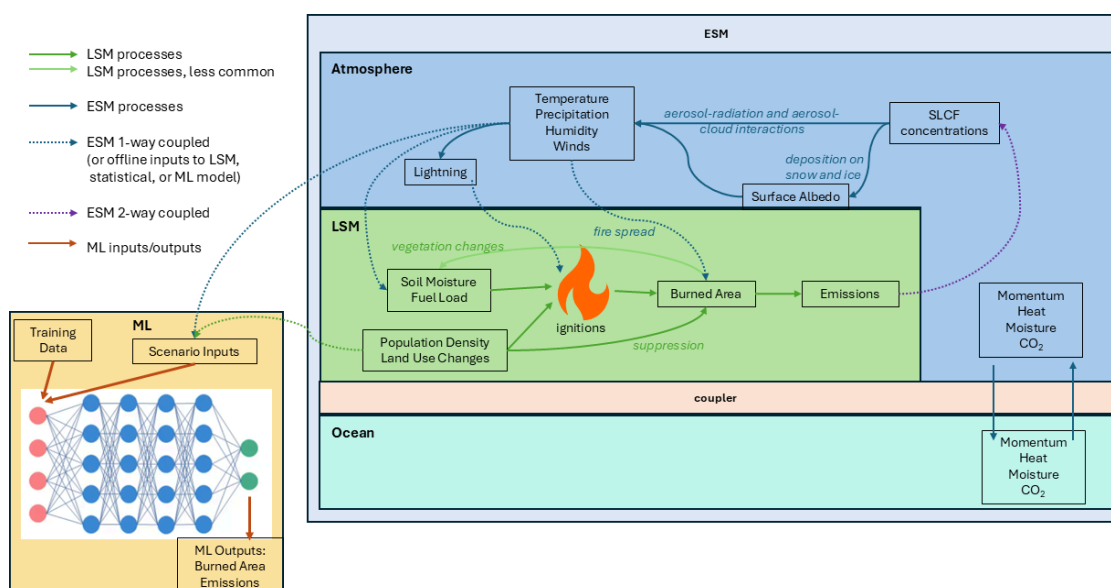
Reviewer #1 (Remarks to the Author):

I read this as a generalist (rather than modelling expert) and found the paper extremely useful and thoroughly researched. It is well written and fills an important niche in the global wildfire literature.

I have not major criticisms and only a few suggestions.

1. The authors talk a lot about various earth system components that interact in various ways, and they have a nice diagram of some of these factors. I think an additional diagram showing linkages would help appreciate interactions and feedback loops, especially if the linkages were colour coded to show relationship that are routinely or rarely considered in LSM and ESM. I appreciate these relationships are indicated in the Figure 1, but reading the text there are numerous climate, chemical and biological variables that are discussed that could be put into an influence diagram to highlight the complexity of global wildfire modelling.

[AR]: Thank you for this suggestion. We have added a new figure (Fig 3 in revised manuscript) to better illustrate the linkages and feedbacks in LSMs, ESMs, and ML models, including pale and dotted arrows to indicate which are less commonly included in models. It now appears in our Sec 3, where we now better synthesize the different modelling approaches (in Sec 3.4) .



2. The authors talk a lot about model predictions and uncertainties, but can they state clear take homes in a table or a text box. What aspects of fire regime on Earth are most likely and least likely to change. It would be impactful to have this information available at a glance.

[AR]: Thank you for this suggestion, and the 2nd reviewer had a similar suggestion. Thus, to address both concerns, we have added a new text box in Sec 4.3 of the revised manuscript to highlight the most important uncertainties, and a new take-home message Figure (Fig 7 in revised manuscript) in Section 5. On key uncertainties:

Source of uncertainty	Type	Relative importance	Why
climate projections	cross-cutting	very high	propagates through every framework
vegetation/fuels	structural	very high	controls fuel availability
human-fire interactions	structural/ parametric	high	dominant tropical and populated regions
fire process representation	structural	high	spread, ignitions, plume height
lightning	structural/ forcing	moderate	important in remote boreal regions
observational constraints	structural	moderate	limits evaluation and ML training
numerical parameter choices	parametric	lower	tuning for historical agreement

A general at-a-glance summary/take-home messages (Fig 7):

☒ prognostic ☒ prescribed ☐ missing/none ☒ data-driven/trained	LSM	Typical ESM (without interactive fires)	ESM with interactive fires	ML/statistical models
Strengths	separates drivers; same model can use different forcing datasets to quantify forcing uncertainty.	coupled climate, well-vetted, simulates transport of prescribed fire emissions	resolves smoke transport and two-way climate-chemistry-ecosystem feedbacks.	extremely fast after training and suitable for large ensembles and rapid projections.
Limitations	cannot simulate smoke transport or two-way atmospheric feedbacks; sensitive to forcing and fire/fuel parameterizations.	feedback and fire emissions not consistent with climate, thus future projections are limited to fire weather and/or scenario data	costly; coupled and compensating biases plus internal variability obscure driver attribution.	only as reliable as the training data; domain shift, spurious relations, and physical inconsistency remain risks
Shared challenges	imperfect burned-area, emissions, fuel, and human-activity data; meteorological forcing bias; scale mismatch and unresolved extremes; sparse validation; structural uncertainty			
Best use cases	vegetation transitions, fire process development	seasonal and century-scale fire weather predictions and projections	global climate feedbacks, century scale projections	seasonal fire danger, operational prediction, uncertainty diagnoses
Components:				
atmospheric dynamics and chemistry	☒	☒	☒	☒
vegetation dynamics	☒	☒	☒	☐
fuels (load, moisture, continuity)	☒	☒	☒	☒
fire processes (ignition, spread, intensity)	☒	☐	☒	☐
emissions (gases, aerosols)	☒	☒	☒	☒
climate feedbacks (radiation, clouds, circulation)	☐	☒	☒	☐

3. The focus of the review is modelling future global fire activity, but can the authors indicate how the models are reinforcing (or counteracting) current empirically derived trends in wildfire activity. If the model predictions are reinforcing current trends this reciprocally strengthens both the inference of current trend analyses and faith in model predictions.

[AR]: Thanks for this suggestion to strengthen our review of the model evaluation of historical trends, as they lend confidence to future projections. We have added more content to section 3 (in the new Section 3.4) to provide details about how well these models perform in the historical time period, and what that means for their usefulness in future projections:

“For all modelling approaches, assessing historical performance is a prerequisite for reliable future projections. However, it is critical that models get the right answer for the right reasons, as the drivers of future trends may differ from those in the historical period, and models are tuned to present-day conditions. For instance, LSMs and ESMs sometimes fail to reproduce the observed decline in global burned area since 1850 and this was attributed to fire suppression and land use changes \cite{arora2005, kloster2010, hantson2020, mezuman2020, li2024, teixeira2025}. However, trends are already changing since the satellite era \cite{zou2019, zhu2022, zhang2023a, kavhu2025, teixeira2025}, and these trends may continue to change in the future, via policies around deforestation and fire management. At the same time, CO₂ fertilization leads to enhanced vegetation growth and fuel load, and unprecedented warming will increase flammability. Representing regional differences in dominant fire drivers remains challenging. Evaluation based only on global metrics is dominated by tropical burning and can obscure poor performance in climatically sensitive regions such as boreal forests \cite{arora2005, arora2018, tian2022, seo2025}.

Despite limitations, process-based and ML models serve as useful frameworks for future fire projections \cite{zhang2023a, li2024}. Modern models extend beyond traditional ESM applications that provide meteorological data for estimating fire risk. They can support climate research and policy by examining the role of fire in future climates \cite{druke2019, mezuman2020, teixeira2021}, including the effects of changing fire emissions and land use management.”

Reviewer #2 (Remarks to the Author):

Summary:

This manuscript presents a comprehensive review of current modelling approaches used to project wildfire activity under climate change, including process-based land surface and Earth system models (LSMs/ESMs), as well as statistical and machine learning (ML) approaches. The authors synthesize recent literature to compare model structures, key drivers, and projected changes across regions and climate scenarios, and highlight major uncertainties and emerging directions, including hybrid modelling frameworks and the integration of AI techniques.

The topic is timely and highly relevant. The manuscript is clearly written, well-structured, and provides a very useful and comprehensive synthesis of a rapidly evolving field.

In my opinion, It will be of clear interest to researchers working on wildfire-climate interactions and Earth system modelling.

[AR]: Thank you for your thorough review and recommendations for improvement.

General assessment:

Overall, this is a strong and thorough review that successfully compiles a large and diverse body of literature. The manuscript is particularly valuable in its impressive breadth of coverage, including extensive supporting material, and in its effort to bridge process-based and data-driven modelling paradigms. However, in its current form, the manuscript remains largely descriptive. While it provides an extensive overview of existing approaches and results, it falls somewhat short of delivering the level of critical synthesis and conceptual advancement that would make it a clear “significant advance” for Communications Earth & Environment. Strengthening the analytical and comparative aspects of the review would substantially enhance its impact. In this sense, the availability of detailed model inventories (e.g. in the Supplementary Information) suggests that a more explicit comparative synthesis across modelling frameworks is both possible and desirable.

[AR]: Thanks for this suggestion. While we had some synthesis in Sections 4.4, 4.6, and throughout Section 5 of the original manuscript, we have now strengthened and added to the synthesis in our revised manuscript, including more comparative aspects:

–We have re-written the first half of Section 4, such that the 3 model types are no longer described in parallel, but rather that their common findings are discussed all together. This also greatly reduced repetition and the word count (which was another requirement for our revision).

–There is also a new Section 3.4 that synthesises the different model approaches, and a new box added to Section 4.3 on the key uncertainties in the revised manuscript.

–We also added a new figure to the beginning of Section 5 that also intercompares the model approaches and provides our recommendations for the best applications of each model type.

We believe these changes have greatly improved the level of critical synthesis and conceptual advancement in this Review.

For example, in the revised Section 4.3 on key uncertainties:

Source of uncertainty	Type	Relative importance	Why
climate projections	foundational	very high	propagates through every framework
vegetation/fuels	structural	very high	controls fuel availability
human-fire interactions	structural	high	dominant tropical and populated regions
fire process representation	structural	high	spread, ignitions, plume height
lightning	structural/forcing	moderate	important in remote boreal regions
observational constraints	structural	moderate	limits evaluation and ML training
numerical parameter choices	parametric	lower	tuning for historical agreement

Beginning of Section 5 on intercomparison and best use cases (new Fig 7):

☒ prognostic ☒ prescribed ☐ missing/none ☒ data-driven/trained	LSM	Typical ESM (without interactive fires)	ESM with interactive fires	ML/statistical models
Strengths	separates drivers; same model can use different forcing datasets to quantify forcing uncertainty.	coupled climate, well-vetted, simulates transport of prescribed fire emissions	resolves smoke transport and two-way climate-chemistry-ecosystem feedbacks.	extremely fast after training and suitable for large ensembles and rapid projections.
Limitations	cannot simulate smoke transport or two-way atmospheric feedbacks; sensitive to forcing and fire/fuel parameterizations.	feedback and fire emissions not consistent with climate, thus future projections are limited to fire weather and/or scenario data	costly; coupled and compensating biases plus internal variability obscure driver attribution.	only as reliable as the training data; domain shift, spurious relations, and physical inconsistency remain risks
Shared challenges	imperfect burned-area, emissions, fuel, and human-activity data; meteorological forcing bias; scale mismatch and unresolved extremes; sparse validation; structural uncertainty			
Best use cases	vegetation transitions, fire process development	seasonal and century-scale fire weather predictions and projections	global climate feedbacks, century scale projections	seasonal fire danger, operational prediction, uncertainty diagnoses
Components:				
atmospheric dynamics and chemistry	☒	☒	☒	☒
vegetation dynamics	☒	☒	☒	☐
fuels (load, moisture, continuity)	☒	☒	☒	☒
fire processes (ignition, spread, intensity)	☒	☐	☒	☐
emissions (gases, aerosols)	☒	☒	☒	☒
climate feedbacks (radiation, clouds, circulation)	☐	☒	☒	☐

A quick note on the detailed model inventories: while ESM and LSM model data are usually provided in public repos, the output from most published ML studies do not provide downloadable future ML projection outputs. Therefore, our synthesis remains literature-based rather than performing new intercomparisons of raw model output.

From the perspective of impacts and decision-relevant assessments, the manuscript would also benefit from clearer guidance on how to interpret and handle these uncertainties. In particular, a more explicit distinction between critical vs. secondary sources of uncertainty, and their implications for projection credibility and usability, would be extremely valuable for both researchers and end users.

[AR]: Thank you. We have expanded discussion of key uncertainties in Section 4.3, where we have added information on which uncertainties are of critical importance, and those that are of less importance, including highlighting them in the new text box in that section (and shown above).

Major comments:

1. Limited critical synthesis across modelling frameworks

A central weakness of the manuscript for Communications Earth & Environment is that it primarily summarizes the literature rather than performing a more critical evaluation. The discussion of LSMs, ESMs, and ML approaches is thorough and informative but largely parallel. In this sense, I am missing a more critical perspective, including a more explicit and direct comparison of their relative strengths, weaknesses, and domains of applicability.

[AR]: We agree that a stronger cross-framework synthesis would substantially improve the paper. Accordingly, we have added a new synthesis section (Section 3.4), and re-written Section 4 such that consistent results across all modelling frameworks are highlighted together and no longer summarized in parallel. A major new take-away from this approach is that despite greatly differing methods, most of the patterns and drivers in the future fire projections from different model types are the same - though highly variable in magnitude*. They also have different, but complimentary best-use cases. Thus, we have added a framework-comparison figure in Section 5 (new Fig 7), and a new conceptual schematic in Section 3 (new Fig 3 in the revised manuscript) – all of which highlight the relative strengths, limitations, and domains of applicability of each modelling approach.

*Beginning of Section 4 now says:

“The results outlined below are broadly consistent qualitative patterns across all model types, despite differences in methodology, though they do differ substantially in the magnitude of projected changes.”

The manuscript does not clearly articulate under which conditions one modelling framework provides more reliable or informative projections than others.

[AR]: We have added additional information about model evaluation in the historical period (in Section 3.4) to help address this and review #1's similar comment. We have also added the new figure that highlights each of their strengths and weaknesses, and suggests the best modelling framework for certain applications (new Fig 7 in Sec 5, and shown above)

In this sense, I encourage the authors to include a more explicit cross-framework synthesis, identifying:

- * where models agree vs. disagree*
- * which sources of uncertainty are most consequential*
- * which approaches currently appear most promising for specific applications (e.g. long-term projections vs. near-term forecasting...)*

[AR]: We understand that our original Section 3, and Sections 4.1-4.3 kept the model framework discussions parallel. That was mainly to set the stage to discuss their unique methods and features. We had Sections 4.4., 4.6 and 5 that discussed in a cross-framework manner. However, in our revision, we have added a cross-framework synthesis to Section 3 (the new Sec 3.4) and have expanded the model evaluation information in Section 3 to address which sources of uncertainty are most consequential (and summarized in a new box in Sec 4.4), and we have re-written the first half of Section 4 (replacing 4.1-4.3 with 4.1 on results of projected changes and 4.2 on drivers of that change) in a cross-framework way, emphasising that the models agree about the direction and spatial pattern of most future fire changes, but vary greatly in magnitude.

We have also added a cross-framework synthesis for the specific applications of:

- (a) Climate-fire interactions/global climate feedbacks
- (b) Informing fire and forestry management (vegetation transitions)
- (c) Informing climate adaptation (e.g. extremes, century scale projections)

via a take-away message figure (new Fig 7) that highlights the strengths and weaknesses of each modelling approach, and their best-use cases for applications.

2. Example of critical analysis: Physical consistency and implications for compound events

Following the last point, an important aspect that deserves more explicit discussion is the difference in physical consistency between process-based and machine learning approaches.

Process-based models (LSMs/ESMs), by construction, enforce internally consistent relationships among key variables (e.g. energy, moisture, vegetation dynamics), ensuring that simulated states remain physically plausible under both present and future conditions. In contrast, statistical and machine learning models, while often highly skillful in reproducing historical patterns, do not guarantee such internal consistency unless explicitly constrained. This distinction becomes particularly important in the context of compound and extreme fire-conducive events, where the joint behaviour of multiple variables (e.g. temperature, humidity, wind, fuel moisture) is critical, particularly under non-stationary conditions and in extrapolation regimes relevant for future climate projections. Models lacking inherent physical constraints may capture correlations in the training data but produce unrealistic or physically inconsistent combinations under novel climate conditions.

While the manuscript mentions limitations of ML approaches in terms of extrapolation and missing feedbacks, a more explicit discussion of physical consistency and its implications for compound events and extreme regimes would strengthen the comparative assessment of modelling frameworks and provide valuable guidance for future developments, particularly in the context of hybrid and physics-informed approaches.

Furthermore, this also reinforces the potential importance of physics-informed machine learning approaches, which may help bridge the gap between predictive skill and physical consistency.

[AR]: Thank you for highlighting and articulating this important distinction. We have revised the manuscript accordingly:

–Our new schematic (Fig 3 in revised manuscript) demonstrates the physical consistency in LSMs and ESMs, and how ML models are different.

–And in our new Section 3.4 of the revised manuscript, we have also added text that explains the internal self-consistency of interactive fire ESMs, and we have added discussion of compound events and the abilities and limitations of ML models to simulate them. For example, part of Section 3.4 now says:

“Statistical and ML models face several important limitations. Their performance depends strongly on the quality and representativeness of training datasets \citep{jain2020, alkhathib2023}. While projections from all model types are sensitive to biases in the underlying climate models, most ML frameworks do not explicitly enforce physical consistency, limiting their ability to simulate coupled Earth system responses \citep{li2023, liu2016, spracklen2009}. These differences become most apparent when simulating compound fire-weather extremes, where fire behaviour depends on the simultaneous evolution of interacting meteorological and fuel variables. As these relationships may change under future climates, ML models trained on historical observations are less reliable when extrapolated beyond the range of past conditions than process-based models that explicitly represent the governing physical processes. Predictive skill also varies geographically, with lower performance often occurring in regions with complex topography or ecological transition zones \citep{balshi2009, illarionova2025}.”

3. Lack of quantitative or structured comparison

Closely related to the previous point, the manuscript would benefit from a more structured and, where possible, quantitative comparison of modelling approaches.

Examples of possible improvements:

** summary tables comparing key features of LSMs, ESMs, and ML frameworks (e.g. feedback representation, spatial scale, computational cost, ability to extrapolate)*

[AR]: We believe our new schematic (Fig 3 in revised manuscript) contains the structured comparison of modelling approaches, as it highlights key features of LSMs, ESMs, and ML frameworks. Also our new Fig 7 might address this as well.

Regarding quantitative comparisons:

–Figure 5 in the original manuscript (which is now Fig 4 in revision) contained quantitative comparisons of future fire projections across model types.

–Figure 6 (in original, now Fig 5 in revision) was also based on quantitative data in the literature. However, since the regions, fire metrics, and scenarios differed quite a bit, we expressed their data all in one figure in a more qualitative manner, since an apples-to-apples quantitative synthesis wasn't possible.

–In the revised Supplement, we have added Table S6, which contains the quantitative data from the modelling approaches that we had gathered from the literature to generate Figures 5 and 6 in the original main paper (Figs 4,5 in revised main paper).

** more systematic comparison of projected changes (ranges, sensitivities) across studies*

[AR]: Our Figure 5 in the original manuscript (now Fig 4) provided a systematic quantitative comparison of projected changes across studies and model types for directly comparable fire metrics (global burned area and global fire emissions). While there are many more papers using these types of models for future fire projections (discussed throughout the paper), they are typically all for different regions, and/or different fire-related variables (e.g. FWI, FFDI, KBDI, etc - see Supplement Note 3 and new Table S6), and thus, can't be directly compared to one another quantitatively. So Figure 5 (now Fig 4) brought together the most commonly reported metrics (burned area and fire emissions) from the most common region (global) across studies that could be directly compared quantitatively.

In terms of ranges, those are represented when the same study appears as more than one bar in Fig 5 (now Fig 4).

Figure 6 (now Fig 5) also brought together the projected changes across regional studies, in as systematic manner as possible, though could not be directly quantitative given the range of data types that those studies reported on.

In the revised Supplement, we have added a large table (new Table S6) that contains all of the quantitative data from future fire projections that we extracted from the literature, with indications on which values were included in Figs 5(now 4) and 6 (now 5).

At present, many statements remain qualitative, making it difficult for the reader to assess the relative robustness of different approaches. In my opinion, the Supplementary Information already compiles detailed model descriptions and classifications; distilling this information into comparative tables or synthesis figures in the main text would substantially improve accessibility and impact.

[AR]: We have added Table S6 to the supplement, which demonstrates the quantitative data that we had gathered from the papers. When possible (apples to apples comparison) we included the quantitative data in our original global synthesis figure - Fig 5 (now 4), and summarized qualitatively the apples to oranges data in the original regional synthesis figure - Fig 6 (now 5). Each modelling framework can make contrasting projections, and there are no systematic uncertainties across models. We have added a new comparative table to Section 5 (new Fig 7), which more clearly let's the reader know the strengths and weaknesses of the different approaches as well as other take-away information.

4. Prioritization of uncertainties, which ones are critical and which ones secondary?

The manuscript identifies a wide range of uncertainties (climate forcing, vegetation dynamics, lightning, human activity, model structure), but these are presented largely as a list.

What I am missing is a stronger effort to prioritize uncertainty sources and to clearly distinguish between dominant vs. secondary drivers and structural vs. "parametric" uncertainties, including an assessment of which uncertainties most strongly limit predictive skill and the credibility of the projected changes. In my opinion, providing such prioritization would significantly improve the manuscript's usefulness and impact.

[AR]: Thank you for this suggestion. We believe the revised Section 4.4 to now contain the prioritization of uncertainties, including this new text and box:

"Future fire projections remain highly uncertain across all modelling frameworks because fire activity depends on interacting climatic, ecological, and anthropogenic processes that are themselves difficult to project under climate change. The dominant uncertainties limiting confidence in future projections are structural rather than parametric. Differences in future climate forcing and in the representation of vegetation, human activity, and fire processes generally contribute more to projection spread than uncertainty in individual parameter values. Parameter uncertainty remains important, particularly for emissions, combustion, and atmospheric chemistry, but typically has less influence on projected long-term changes in fire activity than structural differences among modelling frameworks."

Source of uncertainty	Type	Relative importance	Why
climate projections	cross-cutting	very high	propagates through every framework
vegetation/fuels	structural	very high	controls fuel availability
human-fire interactions	structural/ parametric	high	dominant tropical and populated regions
fire process representation	structural	high	spread, ignitions, plume height
lightning	structural/ forcing	moderate	important in remote boreal regions
observational constraints	structural	moderate	limits evaluation and ML training
numerical parameter choices	parametric	lower	tuning for historical agreement

5. Limited conceptual advance

While the review is comprehensive, it does not clearly introduce a new conceptual framework or synthesis that changes how the reader understands the field. For a journal of this level, it would be desirable to include:

- * a more explicit conceptual synthesis (e.g. a unifying framework linking climate, fuels, and human processes across modelling approaches)*
- * or a clearer articulation of key paradigms or regimes governing wildfire projections*

At present, the manuscript reads more as an extensive and well-executed compilation than as a “transformative” synthesis.

[AR]: We have added a more explicit conceptual synthesis (revisions to Sections 3 and 4, including new Figures (Fig 3 and 7 in the revised manuscript) to bring together the modelling approaches. We have also added clearer articulation of key processes governing wildfire projections (e.g. re-written Section 4.2, and edits to Section 5).

6. Treatment of human-fire interactions

The role of human activity is acknowledged throughout the manuscript, but it is often treated as one of several uncertainties rather than a central driver. Given the strong evidence that land use, suppression, and ignition patterns dominate fire regimes in many regions, the manuscript would benefit from a more integrated treatment of human-fire interactions across all modelling frameworks and a clearer discussion of how different models represent (or fail to represent) these processes in a realistic way.

[AR]: Thanks for this suggestion to highlight the importance of human activity for fires. We have added it to our new box in Section 4.4 in the revised manuscript and include it in our new schematic (Figure 3 in revised manuscript). However, the main focus of this literature review is about the need for models to accurately capture climate change and its feedbacks with wildland fires. There are several reasons why we don't devote more content to the human-fire interactions:

- (a) CMIP6 and CMIP7 only include land-use change and do not include climate change as drivers for future fire emissions. Thus, it is already *only* the human-fire interactions that these large multi-model frameworks include that go on to inform the IPCC and the COPs/climate policies. Our lit review aims to highlight the importance and need for the climate-fire interactions, as that is less commonly done, but has a significant impact on the results (e.g, Fig 6d-f in revised manuscript highlights the difference in emissions when climate is taken into account vs what CMIP6 provides), thus, the critical need remains.
- (b) Most results agree that the human-fire interactions have a large impact and influence in the *tropics* where there is more forest clearing for agriculture, and regular agricultural and savanna burning occur. While the tropics have the most biomass burning emissions globally, they are trending downward (or stable). Whereas, extra-tropical fire regions have less of a human-influence, more climate-sensitivity, and are increasing with climate change.

- (c) We had already included human-fire discussions in Sec 3 (on how different models represent), as well as in Sections 4.4, 4.6, and 5.4 of original manuscript [all integrated (cross-model-framework) treatment of human-fire interaction], and this content remains in the revised manuscript.

Therefore, while both human and climate feedbacks are important, it is our opinion that the bigger gap in modelling wildland fires is in climate feedbacks, rather than human feedbacks, which is why we don't want to devote more content on human feedbacks than we already have in the already-too-long manuscript.

7. Towards actionable projections: Link between model evaluation and projections

The manuscript highlights the need for actionable projections, but the discussion remains relatively general. It would be valuable to more explicitly link model capabilities and uncertainties to decision-relevant metrics and discuss how current limitations affect the usability of projections for fire management and adaptation planning.

[AR]: Thank you for these suggestions. The new box on key uncertainties and Figure with suggested applications addresses this concern related to explicitly linking model capabilities and uncertainties to usability of these projections. There is also additional context about model evaluation of historical trends and implications for future trends added to Section 3.4.

In addition, the manuscript compiles a substantial body of literature on both model evaluation and future projections (as also reflected in the Supplementary Information). However, the connection between model skill in reproducing historical fire behaviour and confidence in future projections remains underdeveloped. A more explicit discussion linking model evaluation, identified biases, and projection credibility would significantly strengthen the synthesis and help readers assess which projections are most trustworthy.

[AR]: Reviewer #1 had a very similar comment, and to address both, we have added more content to section 3 (in the new Section 3.4) to provide details about how well these models perform in the historical time period, and what that means for their usefulness in future projections. The new text states:

“For all modelling approaches, assessing historical performance is a prerequisite for reliable future projections. However, it is critical that models get the right answer for the right reasons, as the drivers of future trends may differ from those in the historical period, and models are tuned to present-day conditions. For instance, LSMs and ESMs sometimes fail to reproduce the observed decline in global burned area since 1850 and this was attributed to fire suppression and land use changes \cite{arora2005, kloster2010, hantson2020, mezuman2020, li2024, teixeira2025}. However, trends are already changing since the satellite era \cite{zou2019, zhu2022, zhang2023a, kavhu2025, teixeira2025}, and these trends may continue to change in the future, via policies around deforestation and fire management. At the same time, CO₂ fertilization leads to enhanced vegetation growth and fuel load, and unprecedented warming will increase flammability. Representing regional differences in dominant fire drivers remains challenging. Evaluation based only on global metrics is dominated by tropical burning and can obscure poor performance in climatically sensitive regions such as boreal forests \cite{arora2005, arora2018, tian2022, seo2025}.

Despite limitations, process-based and ML models serve as useful frameworks for future fire projections \cite{zhang2023a, li2024}. Modern models extend beyond traditional ESM applications that

provide meteorological data for estimating fire risk. They can support climate research and policy by examining the role of fire in future climates \cite{druke2019,mezuman2020,teixeira2021}, including the effects of changing fire emissions and land use management”

Minor comments

Some sections (particularly in the modelling framework overview) could potentially be slightly condensed to improve focus on synthesis.

[AR]: After addressing all of the above comments, we went back and shortened for conciseness throughout the manuscript to approach the 6000-word guideline. The re-write of Section 4 to present the material in a cross-framework manner has helped with that too.

Figures are helpful; however, additional schematic or summary figures (e.g. comparing frameworks or uncertainty sources) could further enhance clarity. However, this is in my view totally secondary.

[AR]: Our three new display items are targeted to provide these schematics and summaries for the frameworks and uncertainty sources. These are the new Fig 3 (pasted below), new Fig 7 (mentioned and pasted above), and the uncertainties text box in sec 4.3 (pasted above in a previous comment) in the revised manuscript. These helped address this and previous concerns from both reviewers.

