Interactive discussion for egusphere-2026-2906

We sincerely thank Referee #2 for the thoughtful review and constructive comments. The previous version did not always distinguish the physical relevance of a process from its explicit representation in the analyzed models, and that the links among drivers, regional uncertainty, and research priorities were insufficiently synthesized. Some related issues have already been considered during the revision for Referee #1, and we have further revised them in response to the present comments, most notably the incorporation of GFED5 and clarification of the FireMIP forcing protocol. We have re-examined those changes considering the present review, corrected remaining factual errors, and made substantive revisions. Compound comments are divided below into smaller points. Reviewer comments are reproduced in blue, followed by our responses and representative revised text. Section references refer to the revised manuscript.

[General comment] In their manuscript 'Reviews and syntheses: Spatiotemporal Dynamics, Drivers, and Uncertainties of Global Wildfire Carbon Dioxide Emissions', Liang et al. integrate multiple datasets, including remotely-sensed and offline and coupled model simulations, to understand different aspects of global wildfire carbon dioxide emissions, including spatial patterns, trends in present and future projections, drivers and associated uncertainties, and finally make suggestions on future research directions.

While I think idea of the paper is interesting and timely, I have many concerns. Given I recommend a rejection based on this, I will only provide my main criticisms:

[Response to general comment] We appreciate the reviewer's recognition of the timeliness of this topic and thank the reviewer for raising these concerns. We revised the manuscript to better link WCE uncertainties with their underlying causes, updated the analysis with GFED5, clarified FireMIP limitations, reduced over-interpretation of spatial patterns, and made the future research priorities more WCE-specific.

[Comment #1] In general, I would describe this manuscript as a literature review rather than a synthesis. I appreciate that this is a massive undertaking, however, I find the different areas the authors cover are descriptive and disconnected: large parts of text for example quantify extensively uncertainties the authors find in custom analyses of different data sources, which I think during the description could have used more direct context of existing literature given this is a review, and then in different parts in similar detail drivers are described and also potential sources of uncertainty. What I'm missing for this to be a useful synthesis however is linking those findings: Based on literature, can the authors come to any conclusions on what driver drives uncertainty where, to which degree, and what would be pathways forward to constrain this uncertainty?

[Response #1] We sincerely thank the reviewer for pointing this out. We structured the identification of uncertainty sources and their impacts into observational, parameter, and structural perspectives. For observational uncertainty, we identified the under-detection of small fires by coarse-resolution products as a dominant driver of observational uncertainty. For parameter uncertainty, fuel consumption, combustion completeness, and CO_2 emission factors are important contributors, particularly in carbon-dense ecosystems. For structural uncertainty, limitations in the representation of fire-process remain a major source of model uncertainty. We also consolidated the corresponding constraint pathways in Section 2.3 so that the diagnosed uncertainties are directly linked to possible approaches for reducing them. In addition, the research priorities in Section 5 were revised to follow from the main uncertainty sources identified in Section 2.3.

Revised text:

Synthesizing across the historical, contemporary, and future periods, uncertainty propagates through three interacting sources (observational, parameter, and structural) whose relative importance varies by region and process rather than admitting a single global percentage decomposition. For the observational dimension, under-detection of small, fragmented fires by coarse-resolution products is the primary driver of contemporary uncertainty (Lou et al., 2023; Shirvani et al., 2023). Mean global fire carbon emissions in GFED5 are 3.40 Pg C yr^{-1} (2002–2022 average), 65% higher than those from GFED4s (van der Werf et al., 2025), and the increases are greatest in fragmented and human-dominated landscapes where coarse-resolution products miss the majority of small fires (Lou et al., 2023). This factor alone explains the majority of inter-product discrepancies in human-dominated regions. For the modelling parameter dimension, concerns the conversion of detected fire activity into CO_2 emissions, including fuel loads, combustion completeness, and CO_2 emission factors, and is most uncertain where burning extends into organic soils and peat. For the structural dimension, deficiencies in fire process representation are the leading source of bias across all periods (Hantson et al., 2020). Across nine FireMIP ensembles, these gaps result in a global annual fire carbon emission range of 0.91–4.75 Pg C yr^{-1} for modern conditions, and the models largely fail to represent interannual variations in burned area or to simulate fire sizes adequately in key regions (Hantson et al., 2020). Furthermore, prevailing fire-enabled Earth system models do not properly account for carbon release from smoldering peat and organic soils (Eckdahl et al., 2026), precluding the reproduction of pre-industrial biomass burning trajectories recorded in ice-core and charcoal proxies, which indicate substantial centennial-scale variations in boreal fire activity over the past two millennia (Zennaro et al., 2014; Rubino et al., *2016). These quantified impacts are not minor biases, but the dominant contributors to current WCE uncertainty, and form the basis for the constraint pathways below.*

Building on the global driver quantification above, we next outline a region- and process-specific hierarchy of WCE uncertainty, rather than a single global uncertainty budget, and the corresponding constraint pathways tailored to each dominant mechanism. In savannas, croplands, and fragmented human-dominated landscapes, missed small and low-intensity fires are a leading observational limitation; the 65% higher mean global fire carbon emissions in GFED5 relative to GFED4s, with still larger regional differences, demonstrates the potential magnitude of this product-structure effect (van der Werf et al., 2025). The most direct constraint is joint high-resolution burned-area and fire-radiative- energy evaluation, with explicit accounting for shared inputs. In tropical forests, peatlands, and boreal organic soils, burned area alone is less diagnostic because fuel consumption, combustion completeness, emission factors, and burn depth control CO2 per unit area; field measurements and atmospheric inversions are therefore higher-priority constraints. In boreal and temperate forests during extreme seasons, model structure associated with fuel preconditioning, multiday spread, severity, and post-fire recovery becomes prominent, requiring event-scale benchmarking rather than climatological means. Before 1901, recycled forcing and proxy interpretation dominate temporal uncertainty, so proxy–model comparison should test the direction and timing of change rather than absolute CO2 flux. In human-dominated regions, uncertainty arises from conflating ignition, suppression, fuel treatment, land conversion, and fragmentation; separate observations of fire starts, escape, response, spread, and final burned area are needed. These evidence streams use different variables and scales, so the current literature does not support a defensible percentage apportionment of total uncertainty across all drivers. We therefore provide numerical degrees only for directly comparable contrasts and otherwise identify the dominant diagnosed mechanism and its corresponding constraint pathway.

[Comment #2] My impression is that in parts the authors need to be more careful about the presentation of data sources and what they mean / what the most up-to-date sources are:

[Response #2] We conducted a source-by-source audit. The three specific issues are addressed separately below.

[Comment #2a] The abstract boldly states ‘paleoenvironmental archives (charcoal and ice-core black carbon)’ are evaluated in the manuscript. Doing a quick search of ‘paleo’ in the manuscript renders 6 results, all of which mention other studies have used those types of data without mentioning any actual conclusions, and therefore don’t offer valuable insights on what can be gained from those paleoenvironmental archives.

[Response #2a] The abstract overstated the role of the proxy evidence in this review. Referee #1 also raised a similar issue. In the revised manuscript, we have added a comparison between FireMIP simulations and sedimentary-charcoal, ice-core CO, and ice-core δ13CH_4 records, and removed the unsupported reference to ice-core black carbon. Section 2.2 now shows that several proxy records indicate changes in pre-1850 biomass-burning that are not reproduced by FireMIP, providing an independent qualitative test of the direction and timing of simulated change. We also clarify that differences in spatial coverage, atmospheric transport, preservation, and proxy sensitivity prevent these records from being converted directly to CO_2 emissions. The abstract has been revised accordingly to describe proxy evidence as an independent qualitative constraint rather than a quantitative WCE reconstruction. (Abstract; Section 2.2)

Revised text:

Abstract. Wildfire carbon dioxide emissions (WCEs) are increasingly recognized as a major and highly uncertain component of the global biogeochemical carbon cycle, reflecting limitations in constraining their historical evolution, future trajectories, and controlling mechanisms. This review synthesizes evidence from satellite-based emission products, fire-enabled model reconstructions and scenario projections to evaluate global WCE dynamics from 1700 to 2100. Historical reconstructions indicate relatively stable global WCEs during 1700–1850, whereas pronounced divergence emerges during 1851–2000 due to differing representations of land-use change and anthropogenic fire processes. Contemporary satellite observations showed a non-significant negative global WCE tendency during 2001–2025, with a trend slope of -0.1 Pg CO_2 yr^{-2} (confidence interval = 95%; $p > 0.05$). Mean annual WCEs over this period were 9.42 ± 0.86 (annual ensemble means ± interannual standard deviation) Pg CO_2 yr^{-1}, with approximately 83% originating from tropical ecosystems, mainly due to extensive burned area and high gross CO_2 emissions per unit burned area. Multi-model projections suggest increases through the 2040s, followed by growing divergence among socioeconomic pathways. Across datasets and models, annual WCE estimates differ by up to 40%, driven by uncertainties in fire detection, fuel consumption, combustion completeness, emission factors, and fire–climate–human interactions. Despite regional heterogeneity, climate variability, particularly ENSO-related hydroclimatic anomalies, is a major driver of interannual WCE variability. Approximately 44% of global vegetated land have experienced increasing extreme wildfire seasons, particularly in northern high-latitude forests where recurrent burning amplifies carbon losses and weakens ecosystem carbon sinks. We identify priorities for *reducing uncertainty through tighter integration of multi-source observations with fire-enabled process models, improved representation of coupled fire–climate–carbon feedbacks and post-fire recovery, and the application of artificial intelligence to better constrain WCE spatiotemporal variability and enhance predictive methods under increasingly nonlinear fire–climate interactions.*

Section 2.2. Since systematic global fire emission monitoring has only been available since 2000, reconstructions of long-term WCEs have relied on an ensemble of fire-enabled models, and paleoenvironmental proxies. A central initiative in this field is the Fire Model Intercomparison Project (FireMIP), which harmonized input datasets and experimental protocols to benchmark model performance against satellite products and paleo-archives (Rabin et al., 2017). Ensemble means from FireMIP have been widely employed to quantify historical trajectories and project future pathways of WCEs under Shared Socioeconomic Pathways (SSPs) (Li et al., 2024; Rabin et al., 2017; Park et al., 2023), revealing a shift from precipitation-dominated fire regimes in the pre-industrial era to anthropogenic-dominated regimes post-1850 with an impending transition to temperature-driven regimes by 2100 (Pechony and Shindell, 2010). Independent proxy syntheses such as sedimentary charcoal (local-to-regional changes) and ice-core records (broader but transport-dependent signals), indicate increasing biomass burning between approximately 1700 and 1850, in contrast to the weak FireMIP ensemble trend (Ferretti et al., 2005; Wang et al., 2010; Marlon et al., 2008, 2016). Their principal value here is as an independent qualitative test of the direction and timing of simulated change. Recently, efforts are increasingly complemented by machine-learning (ML) surrogates to reduce emission uncertainties, particularly in human-dominated landscapes such as agricultural frontiers and urban wildland interfaces (Burton et al., 2019; Li et al., 2019; Rabin et al., 2017; Yu et al., 2022).

[Comment #2b] The authors have not included GFED5, despite it being available before submission of the manuscript. Given GFED5 also provides a significantly different estimate in burned area and therefore presumably fire carbon emissions, this would be an crucial addition to this manuscript.

[Response #2b] GFED5 is essential for an up-to-date assessment of contemporary WCEs. Referee #1 also raised the similar comment. To address this concern, we have incorporated GFED5.1 for 2001–2022 and the GFED5 near-real-time extension for 2023–2025, expanding the contemporary ensemble from four to five fire-emission products. We recalculated the global and regional magnitudes of WHAT, trends, maps, Table 1, Figs. 1–4 and Fig.7. We also corrected our previous description of the GFED5–GFED4s comparison: the 65% difference reported by van der Werf et al. (2025)

refers to mean global wildfire carbon emissions rather than burned area. The revised manuscript now distinguishes this comparison from differences in burned-area estimates and also notes that the five emission products are not fully independent because they share some satellite inputs, calibration information, and methodological components. (Sections 2.1–2.3; Table 1; Figs. 1–4 and 7)

Revised text:

Moving from historical simulations to the satellite era, contemporary satellite-derived fire emission datasets capture consistent magnitudes and interannual variability of WCEs, with pronounced peaks during El Niño–induced droughts and reductions in La Niña years (van der Werf et al., 2017). Between 2001 and 2025, the five widely used datasets (GFED5, GFED4s, QFED, GFAS, and FEER) reveal broadly comparable global averages but exhibit substantial discrepancies in interannual dynamics and long-term trends (Fig. 3). Notably, this mean-level consensus reflects shared methodological roots (e.g., GFAS calibrated to GFED3.1); thus, the narrow range of global averages vastly underestimates true emission uncertainty, as evidenced by the ~65% higher global fire carbon emissions in GFED5 (3.4 Pg C yr^{-1}) relative to GFED4s (van der Werf et al., 2025). Several products suggest a modest decline in global fire emissions during the early 21st century, largely linked to reductions in burned area across savanna regions of Africa and South America (Andela et al., 2017; Zheng et al., 2021), whereas others indicate relatively stable emissions, underscoring their strong dependence on dataset choice and the continuing uncertainties associated with satellite retrievals and emission modeling.

[Comment #2c] If I understand correctly, the authors rely on the FireMIP tier 1 simulations ranging from 1700-1900 for their analysis of historical reconstructions. Again, I'm missing the acknowledgment of the shortcomings of this part of the simulation. The experiment protocol states that the 1901–1920 climate and lightning forcings were recycled for 1700–1900. The data can be used as a first-order estimate of historical fire activity, but the manuscript should be much clearer that these are model simulations rather than proxy-constrained reconstructions, and that recycling the forcing limits the realism of temporal variability.

[Response #2c] The limitations of the pre-1901 FireMIP simulations were not stated clearly enough in the previous version. Referee #1 also raised the similar comment. The revised manuscript now identifies the 1700–2000 trajectories explicitly as FireMIP model simulations rather than proxy-constrained reconstructions. We explain that the 1901–1920 climate and lightning forcings were recycled during 1700–1900, which prevents the simulations from reproducing the actual sequence of interannual, decadal, and centennial climate variability during that period. The paleoenvironmental records are treated separately as independent qualitative evidence for the direction and timing of biomass-burning change, rather than as inputs used to constrain the FireMIP simulations. These limitations are now stated in Section 2.2, Section 2.3, and in the Fig. 3 caption.

Revised text:

Historical WCE uncertainty reflects both FireMIP experimental design and differences among fire-enabled models. The weak pre-1850 FireMIP trend (<0.015% yr-1) is model derived: recycling 1901–1920 climate and lightning forcing during 1700–1900 suppresses the actual sequence of interannual, decadal, and centennial variability (Li et al., 2019; Rabin et al., 2017). Its disagreement with sedimentary-charcoal, ice-core CO, and ice-core δ13CH4 records therefore provides an important structural diagnostic, although proxy differences in spatial coverage, resolution, preservation, and source sensitivity preclude direct conversion to WCEs (Ferretti et al., 2005; Wang et al., 2010; Marlon et al., 2008, 2016). After 1850, intermodel divergence is amplified by inconsistent representations of human–fire interactions. Anthropogenic ignition and suppression are distinct: ignition schemes estimate fire-start counts or probabilities, whereas suppression may alter escape probability, duration, spread, or final burned area; fuel treatment, land conversion, and fragmentation can modify fuels and connectivity without changing ignition count. Models encode these pathways through different process formulations and empirical proxies. In LPJ-GUESS-SIMFIRE-BLAZE, for example, population density affects an empirical annual burned-area estimate, but the model has no explicit ignition-count or spread calculation (Rabin et al., 2017). CLM4.5 represents crop, deforestation, and peat fires, but soil-organic-matter combustion from peat fires was excluded from the FireMIP emission output; most other FireMIP configurations omit peat fire. Configurations that simulate individual fires generally restrict event duration to one day, whereas empirical burned-area models do not simulate event duration. Table 2 now distinguishes explicit process representation, empirical or indirect human influence, and omitted processes.

Historical WCE trajectories for 1700–2000 shown in Fig. 3 are FireMIP model simulations. The weak ensemble trend before 1850 (<0.015% yr-1) is partly a consequence of the recycled climate and lightning forcing and should not be interpreted as an observation-based historical constraint. After 1850, the models diverge because land-use change, population effects, burned area, fuel consumption, combustion completeness, and other fire processes are represented differently among fire modules (van Marle et al., 2017; Li et al., 2019, 2024). LPJ-GUESS-SIMFIRE-BLAZE does not explicitly simulate anthropogenic ignition counts or individual-fire spread. SIMFIRE *estimates annual grid-cell burned area empirically from fire weather, fuel continuity, and population density, after which BLAZE estimates fuel consumption and vegetation mortality (Rabin et al., 2017). Its declining trajectory therefore cannot be attributed to an explicit anthropogenic- ignition mechanism. Process-based fire modules and proxy records show different trajectories, but the proxy–model comparison identifies structural inconsistency rather than validating any model's emission magnitude because the proxies do not provide direct CO2 fluxes.*

[Comment #3] I am also concerned that some of the interpretations suggest that the authors may not have sufficiently distinguished between processes that are physically relevant to wildfire emissions and processes that are explicitly represented in the datasets and models being analysed.

[Response #3] We agree with this comment. We distinguished among three cases throughout the manuscript: (i) explicitly simulated processes, (ii) empirically, indirectly, or imposed through external inputs, and (iii) physically relevant processes that are not represented in the model. The five examples are addressed separately below.

[Comment #3a] The manuscript claims that 'strengthened fire-suppression policies' are accounted for in SSP1-2.6 and SSP2-4.5 without support in the scenario descriptions.

[Response #3a] SSP labels describe socioeconomic and forcing pathways, but do not themselves prescribe strengthened operational fire-suppression policies. We therefore removed this attribution from the manuscript. Scenario differences are now discussed only in relation to the imposed climate and land-use trajectories, the processes represented in each model, and the projection method. Fire suppression is discussed only where its effects on fire occurrence, duration, spread, or final burned area are explicitly represented in a specific model configuration. (Section 2.2; Table 2)

Revised text:

Extending beyond the observational record, future projections of WCEs (2026–2100) draw on fire-enabled DGVMs (e.g., CLM4.5, ModelE2-YIBs), ESMs from CMIP5 and CMIP6, and machine-learning ensembles (Kloster and Lasslop, 2017; Park et al., 2023; Yu et al., 2022). The multi-scenario simulations of our study suggest a near-term increase of 0.3 – 0.8% yr^{-1} through the 2040s driven by extended fire seasons and intensified drought–heatwave extremes, though differences among SSP-based trajectories reflect the varying model representations of fire processes and emission estimation methods. After mid-century, projections diverge: a high-end climate-risk scenario (e.g., SSP5-8.5) amplifies boreal fire–carbon feedbacks, leading to up to 40% greater cumulative carbon loss than under low-emission pathways (e.g., SSP1-2.6)

(McCarty et al., 2020; Turetsky et al., 2015). Permafrost thaw, peat drying, and organic-soil combustion are identified here as omitted-process uncertainties that could amplify high-latitude carbon losses, but many fire-enabled models omit one or more of these processes (Turetsky et al., 2015; Walker et al., 2020; Blackford et al., 2024). We retain SSP5-8.5 as a high-end sensitivity scenario to explore upper-bound WCE responses rather than a likely or representative future pathway. However, standard CMIP6 ScenarioMIP fields used in some projections are concentration driven and prescribe atmospheric CO_2 (Tebaldi et al., 2021); such simulations quantify WCE responses under imposed concentration pathways rather than a closed fire–carbon–climate feedback, which requires emissions-driven experiments with interactively simulated atmospheric CO_2.

[Comment #3b] The manuscript highlights permafrost and peatland feedbacks as a likely source of increasing fire emissions although not many fire-enabled models include these processes.

[Response #3b] This comment highlights an important distinction between physical plausibility and model representation. Permafrost thaw, peat drying, and organic-soil combustion may amplify high-latitude carbon loss, but many models omit the relevant carbon pools, hydrology, burn depth, or combustion processes. We therefore distinguish between their physical relevance and their representation in the models. Permafrost thaw, peat drying, and organic-soil combustion are discussed as potential mechanisms that may affect high-latitude carbon losses, but they are treated as model drivers only in simulations that explicitly represent these processes. Where they are not represented, we treat them as missing-process uncertainty rather than as drivers of the projected changes. Table 2 also clarifies that CLM4.5 represents peat fire but the FireMIP output excludes peat soil-organic-carbon combustion. (Sections 2.2–2.3; Table 2)

Revised text:

Extending beyond the observational record, future projections of WCEs (2026–2100) draw on fire-enabled DGVMs (e.g., CLM4.5, ModelE2-YIBs), ESMs from CMIP5 and CMIP6, and machine-learning ensembles (Kloster and Lasslop, 2017; Park et al., 2023; Yu et al., 2022). The multi-scenario simulations of our study suggest a near-term increase of 0.3 – 0.8% yr^{-1} through the 2040s driven by extended fire seasons and intensified drought–heatwave extremes, though differences among SSP-based trajectories reflect the varying model representations of fire processes and emission estimation methods. After mid-century, projections diverge: a high-end climate-risk scenario (e.g., SSP5-8.5) amplifies boreal fire–carbon feedbacks, leading to up to 40% greater cumulative carbon loss than under low-emission pathways (e.g., SSP1-2.6)

(McCarty et al., 2020; Turetsky et al., 2015). Permafrost thaw, peat drying, and organic-soil combustion are identified here as omitted-process uncertainties that could amplify high-latitude carbon losses, but many fire-enabled models omit one or more of these processes (Turetsky et al., 2015; Walker et al., 2020; Blackford et al., 2024). We retain SSP5-8.5 as a high-end sensitivity scenario to explore upper-bound WCE responses rather than a likely or representative future pathway. However, standard CMIP6 ScenarioMIP fields used in some projections are concentration driven and prescribe atmospheric CO_2 (Tebaldi et al., 2021); such simulations quantify WCE responses under imposed concentration pathways rather than a closed fire–carbon–climate feedback, which requires emissions-driven experiments with interactively simulated atmospheric CO_2.

[Comment #3c] The manuscript mis-identifies LPJ-GUESS-SIMFIRE-BLAZE as a model with explicit anthropogenic ignitions although Rabin et al. (2017) classifies it as a purely empirical model with no concept of ignition count.

[Response #3c] We thank the reviewer for identifying this factual error. Rabin et al. (2017) shows that SIMFIRE estimates annual grid-cell burned area empirically from fire weather, fuel continuity, and population density, rather than explicitly simulating ignition counts or individual-fire spread. BLAZE subsequently estimates fuel consumption and vegetation mortality. We have revised the relevant column heading, the SIMFIRE-BLAZE entries, and the note in Table 2. Population density is now classified as an indirect human influence on empirical burned area rather than as an explicit anthropogenic-ignition scheme. (Table 2)

Revised text:

Notes: Model characteristics refer specifically to the configurations used in the FireMIP historical simulations and should not be assumed to apply to earlier or subsequent versions. "Explicit anthropogenic ignition" denotes a modeled ignition count or ignition probability; a population-density term in an empirical burned-area equation is classified as an indirect human influence, not as an explicit ignition scheme. "Suppression or management" denotes modeled effects on fire occurrence or escape, duration, spread, fuel properties, or final burned area and should not automatically be interpreted as an operational fire-suppression policy. CLM4.5 represents peat fires, but the FireMIP output excludes combustion of peat soil organic carbon. "√": explicitly represented; "P": partially, empirically, or indirectly represented; "×": not represented.

[Comment #3d] The manuscript implicitly oversimplifies anthropogenic suppression to reductions in the number of ignitions.

[Response #3d] This is an important distinction. Ignition, prevention, suppression, fuel treatment, land conversion, and fragmentation act at different stages of the wildfire emission pathway and should not be treated as equivalent processes. Section 3.3 now distinguishes wildfire ignition count or probability from detection, escape, initial attack, containment, duration, spread, and final burned area. It also separates those mechanisms from changes in fuel amount and landscape connectivity. Table 2 has been revised accordingly so that empirical population effects are not interpreted as operational wildfire suppression. (Sections 3.3; Table 2).

Revised text:

Human activities affect several distinct stages of the fire-emission pathway. Intentional and accidental ignitions alter the number, location, and timing of fire starts, while prevention can reduce ignition probability. Suppression operates after detection and may change escape probability, initial attack, containment, duration, spread, and final burned area. Fuel treatment, prescribed burning, land conversion, and landscape fragmentation additionally modify fuel amount and connectivity and are not reducible to ignition number (Andela et al., 2017; Rabin et al., 2017). Models represent these mechanisms inconsistently: some alter fire occurrence, others duration or spread, and empirical models may relate final burned area to population density without simulating ignition or suppression operations. Observational attribution should therefore distinguish cause-of-ignition records, fire response and escape, perimeter growth, fuel treatment, and land-use change. Human ignitions remain important in deforestation and agricultural frontiers and in densely populated landscapes, including North America and Northeast China (Liu et al., 2012; Knorr et al., 2014; Balch et al., 2017), but a decline in burned area or WCEs cannot be attributed automatically to fewer ignitions or stronger suppression without process-specific evidence.

[Comment #3e] If the future projections are based on concentration-driven CMIP6 ScenarioMIP simulations, these cannot capture the full climate-fire-carbon feedback because carbon-cycle changes do not feed back onto prescribed atmospheric CO2 concentrations.

[Response #3e] The reviewer raises an important limitation of concentration-driven ScenarioMIP simulations. Standard concentration-driven ScenarioMIP simulations prescribe atmospheric CO_2 (Tebaldi et al., 2021). They may represent climate effects on wildfire and, depending on the model, fire effects on vegetation, land carbon, or surface properties, but additional wildfire CO_2 emissions cannot alter the prescribed atmospheric trajectory. We now describe these results as WCE responses under imposed concentration pathways rather than as a closed fire–carbon–climate feedback. Figure 3 and Section 2.3 clarify this limitation, and Section 5.2 identifies paired emissions-driven experiments with interactively simulated atmospheric CO_2 as the appropriate approach for evaluating the full feedback. We added Tebaldi et al. (2021) to the references. (Sections 2.2–2.3 and 5.2; Fig. 3)

Revised text:

Extending beyond the observational record, future projections of WCEs (2026–2100) draw on fire-enabled DGVMs (e.g., CLM4.5, ModelE2-YIBs), ESMs from CMIP5 and CMIP6, and machine-learning ensembles (Kloster and Lasslop, 2017; Park et al., 2023; Yu et al., 2022). Multi-scenario simulations of our study suggest a near-term increase of 0.3 – 0.8% yr^{-11} through the 2040s driven by extended fire seasons and intensified drought–heatwave extremes, though differences among SSP-based trajectories reflect the varying model representations of fire processes and emission estimation methods. After mid-century, projections diverge: high-end climate-risk scenario (e.g., SSP5-8.5) amplify boreal fire–carbon feedbacks through permafrost thaw and peatland drying, leading to up to 40% greater cumulative carbon loss than under low-emission pathways (e.g., SSP1-2.6) (McCarty et al., 2020; Turetsky et al., 2015). Permafrost thaw, peat drying, and organic-soil combustion are physically plausible mechanisms that could amplify high-latitude carbon losses, but many fire-enabled models omit one or more of these processes (Turetsky et al., 2015; Walker et al., 2020; Blackford et al., 2024). We retain SSP5-8.5 as a high-end sensitivity scenario to explore upper-bound WCE responses rather than a likely or representative future pathway. However, standard CMIP6 ScenarioMIP fields used in some projections are concentration driven and prescribe atmospheric CO_2 (Tebaldi et al., 2021); such simulations quantify WCE responses under imposed concentration pathways rather than a closed fire–carbon–climate feedback, which requires emissions-driven experiments with interactively simulated atmospheric CO_2.

Historical evaluation should separate protocol-induced behavior from evidence about actual fire history. Future FireMIP-style experiments would benefit from transient pre-1901 climate and lightning forcing to resolve the recycled forcing bias in historical FireMIP simulations, with sedimentary-charcoal and ice-core records used to test the direction, timing, and broad spatial pattern of biomass-burning change rather than absolute CO_2 emissions as proxies cannot be directly converted to WCE fluxes (Rabin et al., 2017; Marlon et al., 2016). Fire-enabled models should report separately whether they represent ignition counts, escaped-fire probability, spread, duration, final burned area, fuel treatment, crop and deforestation fire, peat or organic-soil combustion, and post-fire recovery. Priorities are improved representation of small human-dominated fires, multiday extreme-fire behavior, organic-soil fuel consumption, *and the distinct effects of ignition, suppression, and fragmentation (Andela et al., 2017; Li et al., 2019, 2024). Permafrost thaw and peat drying should be interpreted as simulated drivers only when the relevant carbon pools, hydrology, and combustion processes are represented; otherwise, they are omitted-process uncertainty (Turetsky et al., 2015; Blackford et al., 2024). Experiments must also match the feedback question. Concentration-driven ScenarioMIP simulations estimate fire responses under prescribed atmospheric CO_2, whereas a closed fire–carbon–climate feedback requires emissions-driven simulations with interactively calculated atmospheric CO_2 and fire emissions coupled to land and atmospheric carbon budgets (Tebaldi et al., 2021). Both experiment types should be benchmarked against regional and global extreme seasons using burned area, fuel consumption, duration, severity, and gross CO_2 emissions rather than long-term means alone, aligning with the extreme-fire benchmarking principle.*

[Comment #4] I like the suggested analysis in Figure 6, however find the conclusions the authors make based on it quite speculative. It might be the way the results are presented, but to me the map looks like there is substantial spatial variability that makes it hard to identify the clear patterns they authors appear to find.

[Response #4] We acknowledge that Figure 7 (Figure 6 in the previous version of the manuscript) shows pronounced spatial heterogeneity, which makes it difficult to identify a clear spatial pattern. In the revised manuscript, we explicitly recognize and describe this heterogeneity rather than attempting to generalize across regions.

Specifically, we state that the drivers of recent extreme wildfire seasons (EWS) changes, including climatic extremes, fuel loading and continuity, vegetation composition, topography, and human activities, are highly spatially heterogeneous. Moreover, our understanding of these drivers and how they jointly influence recent EWS changes in different regions remains incomplete and poorly constrained by observations. Therefore, any claim of a universal or cross-region pattern would indeed be speculative.

To make the hotspot analysis more objective and reproducible, we revised the criteria used to identify increasing and decreasing hotspots. The caption of Figure 7 now specifies the DBSCAN settings (150 km search radius, min_samples = 2). Increasing hotspots are defined as contiguous clusters larger than 200,000 km^2 with a mean pixel-level change > 0.5 EWS yr^{-1}, whereas decreasing hotspots meet the same area criterion but have a mean change < -0.5 EWS yr^{-1}. We also removed statements implying cross-regional comparability or a universal spatial pattern and now limit the interpretation to the hotspots identified by these criteria.

We retain the inherent spatial heterogeneity in Figure 7, and explicitly attribute this heterogeneity to the spatially variable and still poorly understood factors that influence extreme fire risk. This revision reinforces the core conclusion of Section 2.3 that changes in extreme wildfire risk are strongly region dependent. Consequently, Figure 7 provides an empirical basis for future region-specific research on the mechanisms and drivers of changing extreme wildfire risk.

Revised text:

Mounting evidence indicates that the frequency of the largest and most destructive EWS have escalated worldwide (Zheng et al., 2023; Cunningham et al., 2024; Bowman et al., 2017), further destabilizing the terrestrial carbon balance and amplifying the risk of carbon-cycle feedbacks to the climate system (Jones et al., 2022; Senande-Rivera et al., 2022; Walker et al., 2019). The 2021 boreal fires produced a CO_2 emission anomaly of approximately 1.06 Pg CO_2 above the 2000–2020 mean, larger than the anomalies associated with major recent fire events in Indonesia, Brazil, Australia, and the United States (Zheng et al., 2023). Elevated GHG emissions during EWS can amplify warming, which further leads to more droughts and high temperature, contributing to occurrence and spread of extreme wildfires (Zheng et al., 2023; Senande-Rivera et al., 2022; Jones et al., 2024b). Under the operational one-standard-deviation definition, approximately 44% of the vegetated land showed an increasing frequency of EWS globally (Fig. 7), coinciding with a 54% rise in extreme fire weather days (FWI 95th percentile) (Jones et al., 2022). However, the map in Fig. 7 exhibits pronounced spatial heterogeneity: increasing and decreasing changes coexist and are unevenly distributed, and we acknowledge that this inherent heterogeneity makes a single clear pattern difficult to identify. This heterogeneity is expected, because the drivers of extreme fire risk—including climatic extremes, fuel load and continuity, vegetation composition, topography, and human activities—are themselves highly spatially heterogeneous, and how these drivers jointly modulate recent EWS changes in different regions remains incompletely understood and poorly constrained by observations.

Using a density-based spatial clustering algorithm (DBSCAN), we identified increases in EWS across many regions, with the exception of five negative hotspots (blue boxes) that represent areas of decreasing EWS. One increasing hotspot (red pixels in Fig. 7) is located in high-latitude boreal forests, where the detected increase is consistent with a longer, warmer, and drier fire season. Given the region-specific and incompletely understood drivers noted above, we treat this as a local hotspot characteristic rather than evidence of a general high-latitude expansion applicable to other regions. Outside these objectively defined hotspots, EWS changes are weaker and *highly variable, and are not interpreted as a coherent global trend. A comparison with the hotspot analysis of WCEs (Fig. 2), which revealed three negative hotspots, shows an intriguing discrepancy in the sub-Saharan Africa with an insignificant increase in EWS frequency occurring despite a decline in WCEs. A plausible, though not definitive, explanation is the predominance of low-impact fires associated with declining human ignitions and the historically limited documentation of extreme wildfire trends (Andela et al., 2017; Ramo et al., 2021). These findings suggest that the traditional focus on average fire intensities may obscure opposing patterns in extreme wildfire seasons at the region-specific level, which cause the greatest damage and release the largest amounts of CO_2 emissions (Cunningham et al., 2024).*

[Comment #5] The recommendations for future research directions remain very general. Many of the proposed future directions would apply broadly to almost any Earth-system process. I would expect a synthesis focused specifically on wildfire carbon emissions to identify more concrete, fire-specific research priorities emerging from the uncertainties diagnosed in this study.

[Response #5] To avoid the overgeneralization in our prior outlook, we have thoroughly restructured Section 5 to focus exclusively on WCE-specific priorities, directly addressing the diagnostic gaps in Section 2.3 while removing all generic Earth-system recommendations. The revised outlook is organized into three priorities. For observational constraints, we focus on closing gaps in the WCE estimation chain, specifically addressing small-fire omissions identified via GFED5 and the confusion between gross emissions and net carbon balance. For model improvements in model structure, we prioritize efforts to address the process-representation gaps identified in our FireMIP assessment. This includes using transient pre-1901 forcing, improving the representation of multi-day extreme wildfire behavior (e.g., 2023 Canadian wildfire season), and advocating for emission-driven experiments to capture the full fire–carbon–climate feedback. For extreme events, we prioritize region-specific, end-to-end studies of extreme wildfire seasons. We emphasize the use of extreme season totals and wildfire intensity as key benchmarking metrics rather than relying only on long-term trends. All revised priorities are exclusively derived from the WCE-specific uncertainties diagnosed in this synthesis, eliminating the generic generalizations noted in the previous version. These priorities are fully consistent with the empirical updates in Section 2.3 and the extreme wildfire analysis in Figure 7.

Revised text:

The synthesis identifies three interacting sources of WCE uncertainty: observational limits in detecting fire activity, parameter uncertainty in converting fire *activity into CO_2 emissions, and structural uncertainty in representing fire–climate–vegetation–human interactions. These findings lead to three fire-emission-specific priorities. First, each component of the WCE estimation chain should be constrained with observations sufficiently independent to expose shared product biases, addressing the WCE underestimation in coarse-resolution products relative to GFED5. Second, historical and future models should be evaluated for the fire processes they actually represent and with experiments appropriate to the feedback being investigated, resolving structural biases in FireMIP and CMIP6 models. Third, gross fire CO_2 emissions should be evaluated together with post-fire uptake, reburning, and ecosystem transitions to determine longer-term carbon consequences, distinguishing gross WCE from net carbon balance. Extreme wildfire seasons provide a cross-cutting benchmark because they reveal observational and structural deficiencies that may remain hidden in climatological means, and directly reflect the region-specific heterogeneity.*

5.1 Constraining the WCE estimation chain

Contemporary WCE estimates should be evaluated at the level of burned area, fuel load, combustion completeness, and CO_2 emission factor rather than only through agreement in total emissions among products with shared inputs or parameterizations. Priority gaps include small and fragmented fires in savannas, croplands, and deforestation frontiers (the dominant driver of the 65% WCE underestimation in coarse-resolution products); fuel consumption and combustion completeness in tropical forests, peatlands, and boreal organic soils (where burned area alone cannot explain unit-area CO_2 emissions); and biome- and combustion-phase-specific CO_2 emission factors (Roteta et al., 2019; Ramo et al., 2021; Andreae, 2019; van Wees et al., 2022). High-resolution optical burned area and fire-radiative-energy observations should be combined with field measurements and atmospheric inversions, with uncertainty reported separately by biome and fire type, prioritizing peatlands and tropical forests where inversion constraints are most needed. Atmospheric inversions provide a more independent top-down constraint, but their transport, prior, and observational uncertainties must also be quantified (Chevallier et al., 2019). Measurable evaluation targets should include detection omission and commission, fuel consumption per unit area, CO_2 per unit burned area, extreme-season totals, and cross-product error covariance.

To address widespread confusion between gross WCE and net carbon balance emphasized in Section 2.1, all observational studies must report both total WCE and post-fire carbon uptake synchronously. Machine-learning methods may assist gap filling, data fusion, and diagnosis of nonlinear relationships, but their value for WCE estimation depends on validation against withheld and methodologically independent observations. They should be evaluated for specific fire types and extreme seasons, *constrained by mass-balance and ecological relationships, and accompanied by explicit attribution of uncertainty to observations, parameters, and model structure—and must not replace WCE-specific parameterizations of smoldering peat combustion or fuel moisture dynamics, processes unique to fire research—rather than treated as a substitute for process understanding (Jain et al., 2020)*

5.2 Testing missing fire processes with fit-for-purpose experiments

Historical evaluation should separate protocol-induced behavior from evidence about actual fire history. Future FireMIP-style experiments would benefit from transient pre-1901 climate and lightning forcing to resolve the recycled forcing bias in historical FireMIP simulations, with sedimentary-charcoal and ice-core records used to test the direction, timing, and broad spatial pattern of biomass-burning change rather than absolute CO_2 emissions as proxies cannot be directly converted to WCE fluxes (Rabin et al., 2017; Marlon et al., 2016). Fire-enabled models should report separately whether they represent ignition counts, escaped-fire probability, spread, duration, final burned area, fuel treatment, crop and deforestation fire, peat or organic-soil combustion, and post-fire recovery. Priorities are improved representation of small human-dominated fires, multiday extreme-fire behavior, organic-soil fuel consumption, and the distinct effects of ignition, suppression, and fragmentation (Andela et al., 2017; Li et al., 2019, 2024). Permafrost thaw and peat drying should be interpreted as simulated drivers only when the relevant carbon pools, hydrology, and combustion processes are represented; otherwise, they are omitted-process uncertainty (Turetsky et al., 2015; Blackford et al., 2024). Experiments must also match the feedback question. Concentration-driven ScenarioMIP simulations estimate fire responses under prescribed atmospheric CO_2, whereas a closed fire–carbon–climate feedback requires emissions-driven simulations with interactively calculated atmospheric CO_2 and fire emissions coupled to land and atmospheric carbon budgets (Tebaldi et al., 2021). Both experiment types should be benchmarked against regional and global extreme seasons using burned area, fuel consumption, duration, severity, and gross CO_2 emissions rather than long-term means alone, aligning with the extreme-fire benchmarking principle.

5.3 Linking gross fire emissions to recovery and long-term carbon consequences

Because WCEs quantify gross combustion emissions, their longer-term climate significance depends on the rate and completeness of post-fire carbon recovery. Observational and modeling studies should report gross WCEs together with vegetation and soil-carbon uptake, time to recovery, reburning interval, and evidence of persistent ecosystem transition. This is especially important in carbon-dense tropical forests, *peatlands, and boreal forests, where severe or recurrent fire can shorten recovery intervals, expose organic soils, and shift vegetation toward lower-carbon states, as seen in the boreal EWS hotspots mapped in Fig. 7 (van der Werf et al., 2010; Walker et al., 2019; Rogers et al., 2020; Steel et al., 2021). Coordinated field measurements, repeated optical and lidar observations, atmospheric constraints, and vegetation-model experiments should test whether post-fire uptake offsets gross emissions over ecosystem-relevant timescales and how that balance changes under repeated extreme wildfire seasons, prioritizing region-specific analysis of the unique sub-Saharan Africa discrepancy identified in Fig. 7 where WCE declines despite stable extreme fire frequency. Core metrics should include cumulative gross emissions, recovery half-time, net carbon loss at fixed post-fire intervals, reburn probability, and transition to a lower-carbon vegetation state, which align with the extreme-season total emissions and fire intensity metrics emphasized in Fig. 7.d text.*