

Wildfires review

Authors have produced a remarkable assembly and collation effort! Compliments to authors. Good fit to ESSD. I will definitely recommend publication. I hope authors take some useful lessons from this review?

I understand and support this product as first-of-its-kind, a worthy initial effort intended to guide and stimulate progress on fire research. Again, good on the authors for great start! With second version promised, authors can claim 'good idea, address it in next version'. Much of what follows they will want to fix in first version! Authors and journal want best-foot-forward with this product! But some changes could delay?

I read this multiple times! Unfortunately, I feel that I understand perhaps less than half of what authors present. Why should I understand; I do not and have not engaged in wildfire research? But, as former Director of World Climate Research Programme and as co-founder and - for a too-long time - sole chief editor of this journal, I should represent at minimum a partially-competent somewhat-knowledgeable reader? Authors will of course, following scientific instincts, attempt to correct my errors and mis-interpretations. But, please take a larger view: of me as interested supportive (somewhat) knowledgeable reader. If authors have not reached and convinced me, what have they done wrong? How do they need to change their messages? If I can't read this, how can authors expect fire managers (I deal with many, on local as well as forest-wide issues) to extract anything useful?

The manuscript remains way too long! (I deal with length issues below.)

For research audiences or (as authors hope) for "policymakers, disaster management services, firefighting agencies, and land managers" (line 131), this manuscript contains far too many deficiencies. Language ('extremes', 'serious', 'emergency', intense') remains vague and, therefore, difficult. Too many acronyms remain undefined. Maps and figures remain unreadable; figure legends in some case prove wrong. Dates of coverage and extents of observations remain inconsistent. This reader needs to scroll upward (backwards) 10s of pages and 1000s of words to find methodological details. A reader learns that satellite measurements of BA, focused necessarily on larger fires, might miss as much as 90% of burned areas (line 376; roughly 4 million km² might actually total more like 7 million?) while, in later sections, same reader confronts author discussion of -0.2% BA anomalies? In haste to meet deadlines, authors have neglected chances and needs to check their own work and to, necessarily, impose collective oversight? To repeat: they make a great start! But they need to better guide readers and data users on their choices, their assumptions, their outcomes, and their conclusions.

Language: Authors take a strange but perhaps necessary approach: apply confusing ill-defined adjectives first, then later (Section 4.2.1, line 3085) highlight uncertainties and needs for improved definition. Unless authors want to dictate existing or develop new terms, I see no good alternatives to present approach. For this reader, knowledgeable disclaimers by these authors (shortened by 50%?) should have appeared earlier rather than later.

Uncertainties: This reader understands that wildfires represent one of our most-complex challenges. How do they start? How fast and how much do they burn? How much do fuel load, soil moisture and weather (large-scale or self-generated) influence specific fires. How will observations, predictive skills, and attributions change in unknown future? Our basic task remains identifying and then working to reduce uncertainty. To then read about 99% confidence of -0.2% changes in BA seems, frankly, absurd. I don't doubt that one can make such statements with statistical certainty. But, should we? We know that preferred fire observations carry uncertainties near $\pm 100\%$. We know that even our best GCMs carry very large uncertainties, particularly for next decade or one after that. We know that reanalyses, an outcome of same systems we use to specify error matrices, likewise carry substantial uncertainties. Fire models, based on these uncertain sources, very likely amplify uncertainty; at minimum they combine and must depend on all underlying uncertainty.

ESSD guidelines specify "Explicit uncertainty accounting and analysis". Here authors push those boundaries. Readers will not doubt techniques nor skills! They will, however, as I do, want to see authors' rational and best estimates to work within substantial uncertainties. Where authors need to demonstrate valid tools for handling small differences in BA, we need to know that those authors at the same time appreciate the difficulties and larger uncertainties of this science. I recommend the authors introduce an uncertainty budget early in manuscript? Show and admit where uncertainties exist? Outline your approaches: ignore, compare, test, apply two different models, whatever. Show us, however, that you understand the larger questions of uncertainties even when you must proceed.

Atmospheric Climate Modes At least a few of these authors adopt mode (ENSO, IOD, POD, etc.) labels and literature. (Please remember that a long time ago, before birth dates of some of you, I led a big year-long field program focused on western tropical Pacific and ENSO.) I recommend that this manuscript and this effort avoid such references and discussions. For two reasons. A) This manuscript really does not need such discussions to make basic points about S2S predictability? Your results prove, in this case, no better nor no worse - but entirely dependent on - S2S forecast skills. Difficult to dismal, at least at certain time scales, so far? B) Whatever the mode, it depends on an original identification and description now at least 20 (perhaps 40) years old. Those modes, which

initially had at-best marginal utility, have proved increasingly unstable and unreliable. These authors really do not want to step into vigorous ongoing scientific debate, nor have work judged by changing, perhaps eventually no longer valid, mode descriptions?

Length: too long. For many reasons.

1. Challenging for journal staff to process (typeset, proofread, etc.). We (ESSD) depend on their expertise! We should not abuse.
2. Difficult for readers. Results 10s of pages after Methods. No hot links, not easy to move around. Many apparent redundancies.
3. Impassible for fire managers. They will not use this! They will dismiss as too long, too obscure, too full of technical jargon.

Even if authors decide this first version targets other fire researchers rather than 'stakeholders' (fire & resource managers, general public), they still need to understand and respect reasons 1) and 2). I recommend a target of 80+5 pages! Even accepting necessary additions (Table of Contents, for example), authors can easily meet my 80-page target by moving large sections to Supplement, clarifying and sharpening remaining text, and - starting from a lean short product - accommodate eventual inevitable additions. You want this product to grow into respected forum? You will need, in that case, to accept substantial additional ideas and contributions!

Assumptions for 80 pages:

- Move detailed continent-by-continent descriptions to Supplement. I know authors put a lot of work into this, but it offers very limited useful information to overall manuscript goals. Saves: 12 to 13 pages (~600 lines) plus associated figures and reference?
- Delete most discussions of ENSO, IOD, etc. Saves: 2 to 3 (estimated) pages, plus a few references?
- Keep 3.1.1 Highlights OR Section 4.1 Summary but not both. This reader would prefer Section 4.1 but moved near top of manuscript. Saves 1 (estimated) page plus a figure or two?
- Eliminate discussion and outputs from PoF because, in the end, that tool contributes very little to important outcomes? Saves perhaps 1 page?
- Check and then use acronyms correctly and consistently. Saves 100 (or more) words?
- For several sections, authors present highlight section followed by one paragraph details section. Easy to combine / reduce these in some cases?
- Reduce redundancy. Perhaps by adopting, for each major effort (descriptions, predictions, projections, etc.) an organization of methods

first followed directly by outcomes, in sequence. I predict we might gain at least one page per section, mostly by deleting redundancies?

- Rewrite entire manuscript focusing on brevity and clarity.

I rewrote sections to test this last suggestion (my supplement). I achieved length reductions (by word counts) of 25%, 20%, 28% and 12% in these four small samples; I believe expecting 20% reduction seems reasonable. 20% * nearly 90 changeable pages saves another 18 pages? I might, for a few paragraphs or news article, assign differences between my wording and authors' wording to personal language preferences? In this case however, with length a primary problem, authors need to make substantial text reductions!

Necessary additions: Table of Contents, with hot links! One page.
Acronym list (could go in Appendix). Two pages.

By rough accumulation of assumed saving plus 3 additional pages, one easily gets to 80 net (working text) pages?

Understanding this as initial foray, I sort recommendations into three sections: now (mostly proof-type comments that authors will want to correct to make this version 'presentable'); now or future (as mostly related to length or clarity) entirely at authors' discretion; and future (for authors' discussion as they move forward).

Note: I use terms 'reader' and 'user' interchangeably.

Now

Line 108: "This" what? "Driven"? "Dampened"? This BA vs CO2 emissions pattern? This emission pattern? You do not want to allow uncertainty on the part of readers. 'Clearer than the truth', particularly in abstract.

Line 118: "fuel load and direct human suppression often modulated areas with anomalous burned area"? 'All' areas? 'Other' areas? Areas other than Canada and Greece?

Line 121: "extreme events". Again, use of ill-defined term 'extremes'. 'Extreme events' could include landslides, floods, etc. You want to keep reader's focus on wildfires?

Line 128: "extreme anomalies" which differ from "moderate positive anomalies" by how? Explanations, including statistical certainties / uncertainties offered later but here (again) readers will need to find exact language. So many (too many) readers will glance at abstract to check whether they want to peruse further.

Line 135, Short Summary: Excellent. ESSD editors might consider requiring similar short summaries of all papers?

Line 151: again "extreme fires", use of a fraught descriptor that authors will discuss (and, disparage?) later. I will stop recording these instances. Basic question: if authors know they confront an ill-defined (at best) term, should they use it (so often) themselves?

Lines 160 to 164: Authors should read carefully. To this reader, closing sentence lacks a conjunction ("and") or something. Confusing as written?

Line 180: "is significant" Statistically significant? Generically significant? Worrisome? (Authors used term 'concerning' in previous sentence.)

Line 193: very valid point but many ESSD readers will not know 'infamous' Wall Street (NYC) image? Slight alteration or clean up of closing sentence?

Lines 196 and following: Authors recite valid points about C emission but, as they well know, atmospheric CO₂ represents residual between emissions and sink. Point here: wildfires count as land emissions, no longer a land sink? Clean up and perhaps shorten particularly introductory sentence for this paragraph?

Line 205, 206: "undermine the regenerative capacity of forests (Nolan et al., 2021a) and the habitats of many endemic species being degraded in biodiversity hotspots (Ward et al., 2020). 'Undermine' regenerative capacity and habitats? With resulting degradation? Awkward sentence, needs clarification.

Line 207: Authors report correctly changed assumptions about post-fire regrowth but sink or source implications of "fire-C cycle" have always proven difficult? Perhaps a clarification or certification of one term while net equation remains unclear? Important point but needs attention to language.

Line 209: "degraded and transformed" Wildfire may degrade and transform any land, regardless of ownership or occupation. Impacts may prove more challenging or detrimental for native or Indigenous people but that represents a function of ownership, not of fire intensity? Authors allow confusion here between "fire types" (Line 211) and community impact? Valid points but deserve better itemization?

Line 215: "policymakers and involve coordination with many other stakeholders" 'that involve'? Or, 'that require'? We suffer too often when policymakers act in the abstract, without consultation with stakeholders?

Line 232: "seen as a key tool for achieving Net Zero" This reader does not see C

offsets as key nor useful tool; nor will other ESSD readers. Perhaps 'used' or 'applied'?

Line 235 & paragraph following: Very good questions. Can we expect 'stakeholders' to wade through 124 pages to find tools or answers?

Line 242 & paragraph following: Admirable sentiments but authors could reduce this paragraph by 1/2? Again, in response to final sentence about relevance: relevant perhaps but not easily available?

Line 260: Sentence here (starting from "We incorporate ...") needs change in punctuation or in use of conjunctions to clarify exact intent? Later, why do authors, here and throughout, capitalize 'Earth Observations'. Do they mean to indicate satellite observations? Something different or special?

Line 292: Lose first sentence; not needed? Start with 'We identify' ... in second sentence?

Line 302 "Burton, Lampe, et al., 2023": Why does this particular citation show up, throughout the document, with two names? From citation itself (Burton et al. 2023) it looks like standard reference? Something in bibliometric software? Something the authors wish to highlight?

Line 348 and following: Please confirm all dates / durations. How, for example, can this manuscript report data to Feb 2024 if BA updates only go to Andela and Jones (2023). Andela & Jones in reference list actually shows as 2024?

Please also check and confirm all acronyms in this section. This reader knows (but many will not) NASA, MODIS, ECMWF, CAMS, GFAS, GFED, etc. Consistency first then shorten where possible. Careful consistent use of these acronyms (e.g. GFAS) could eliminate many subsequent words?

Line 373 & paragraph following: important points about fire size limitations which emerge multiple times later but, here, authors could reduce this section by 20%?

Line 415, Table 1: Continent (row 2) shapefiles exist in all GIS software, open (e.g QGIS) or proprietary (e.g. ArcGIS). Some readers will operate under institutional ArcGIS licenses but many will rely instead on open GIS software. Make this source generic rather than specific?

In my version, wrapping (at page 8) causes disconnect problem in RECCAP2 line, e.g. Ciais et al? Authors will know what they intend ...

Line 425: Authors use different list numbering here (capitalized Roman numerals)

than in earlier text lists (e.g at lines 274 and following where they used Arabic numerals). Copernicus (publisher) no doubt recommends and adheres to standardized approach?

Line 430: ", and."? Superfluous or something missing?

Line 446 & paragraph following: Good important stuff, readers will need and want this info, but - again - authors could help readers (and publisher) but reduction entire section by $\geq 10\%$?

Line 456, "Global Fire Atlas": Why use extended title when you have already defined acronym? Shorten in all ways possible!

Line 471 & paragraph following: Useful, necessary information but authors could convey identical information in 50% fewer words?

Line 524, Table 2: Again, potentially a wrap problem around page 10 to 11, but Table 2 in my version contains duplicate entries for the Africa team? E.g. both on page 10 and again on page 11, duplicate lines?

Unfortunately, I need a map to understand section 2.3 and following.

FWI, a real-time fire forecasting tool, used to assess how 'early' (how much in advance of actual fire) users could predict that particular fire. FWI requires as inputs weather, dead fuel moisture (only dead, not living above ground?) Predicts ignition possibilities and spread rates. One advantage: moisture over three levels (depths) of soil?

Then, to address longer time scales for prediction, these authors look at FWI-compatible data (conditions) but from longer-time prediction systems (e.g. by ECMWF)?

To better understand detailed predictors (beyond weather and moisture as used by FWI), authors evaluate two additional models (applied only to their focal events?): PoF and ConFire. Here they need to introduce concept of Active Fire (AF) to supplement and complement BA? Because PoF predicts only AF? As aside, note relative incompatibility of AF with BA?

Drivers group into categories (= controls?) but drivers certainly interact (weather & moisture) and (?) predictors = drivers = single variables? Reflect interactive drivers by including them in more than one category/control?

PoF and ConFire require weather, fuel abundance, and fuel moisture, plus other? For 'other' PoF relies on vegetation type (forest vs grassland?), urban fraction and

ography, ConFire for 'other' uses land use type, urban vs rural populations, and lightning (ignition) from Table 3.

Then reader confronts / needs details of PoF and ConFire. Paragraphs, but easily shortened with no loss of info!

Finally, to develop attributions (particularly but not exclusively to climate change factors), we need to hold fire model constant (use FWI) while imposing climate scenarios (via SSP) according to definitions and recipes developed by IPCC and ISIMIP? So long as outputs of those climate scenarios meet needs of fire model (FWI)? Users need to understand that attribution works on different temporal and spatial scales?

Attribution work undoubtedly needs almost this level of detail: define terms, describe and validate climate models, set up forcing scenarios (following ISIMIP, with another Table?), describe experiments (All, Natural, climate, socio-economic, factual, counterfactual) and replicates, etc. But, do authors need to expend so much text on these processes??

Seasonal outlooks, which remain elusive for many features in many regions but good on these authors for pushing fire! Long but (in this readers view) useless discussion of ENSO, IDO, AMO, PDO, anyone's additional 'natural' mode or oscillation. Useless because best these authors can derive remains 'associations' (lines around 2680). ENSO prediction remains a fraught and failing endeavor itself, with teleconnections (on precipitation, drought, moisture) sometimes positive, sometimes negative, and occasionally, indecipherable. Those imagined non-stationary oscillations emerge only over large regions based on monthly (at best) and ocean-wide averages. These authors should not accept prior ENSO nonsense! Have the courage to show, from their data, no reliable valid correlations, positive or negative? I would start section 2.5 at line 1008: "To look more closely at impacts that historical climate oscillations (eg. ENSO, IOD, others) might have on landscape flammability ...

End of diatribe. Do I have this correct? If I need a map, or graphical equivalent of a map, users will likewise?

Line 527, Bespoke Air Quality: not a main focus, move to Supplement? Or, better fit later in AQ sections/discussion?

Line 545, 2.2 Shortlisting: Good section, authors could write same content with 50% fewer words but no loss of info?

Line 589, "June 2024), and the Canadian Wildland Fire": need one fewer 'ands'

and possible punctuation change to fix this sentence. Question: wouldn't we expect Canadian Wildland Fire Information System to use FWI?

Line 596, "found to correlate": an equal number of papers show no correlations or no consistent correlations? For other regions or reasons or at other seasons, but still? At line 599, I doubt that on rigorous statistical basis, one can document any consistent ENSO etc. teleconnection to fire occurrence or intensity predictability. If authors know differently they need to cite sources? Aragao & Turbo papers do not prove what these authors hope? Authors could reduce this paragraph by ~15% with no loss of info?

Line 612, "seasonal skill is limited to 2-3 months" at best! For specific regions, specific seasons, and specific conditions!

Line 654, "Fire Weather Index": you already introduced acronym FWI. Shorten, shorten, shorten!

Line 663: I know "ESA CCI" as ESA Climate Change Initiative (I used to chair their advisory board) but users may not? Definition not until line 680?

Line 690, "that area": refers only to Greece or to all three focal events?

Line 694, Table 3: In header role, PoF control vs ConFire controls (plural)? Here, authors describe drivers as "individual explanatory variables???" I know but many readers will not: SMOS?

Line 719, "monthly daily means": Confusing, not sure what 'monthly daily' means? Same line "FLAME"?

Line 769, "apply different modelling techniques"; use a word other than 'different' here. In this context, different can imply different from previous, not what authors intend? Paragraph that follows could reduce by 20%?

Line 857: good point about resolution of Greek data but redundant with much of prior paragraphs?

Line 861, "MODIS MCD64A1"; by now reader has seen MODIS product designations so many times, he/she no longer knows which label associates with which product? Develop (or copy) and apply standard abbreviations or acronyms?

Lines 869,870, more jargon: "MODIS Vegetation Continuous Fields collection 6.1 remote sensed data for <60°N DiMiceli et al. (2022) and collection 6 for <60°N DiMiceli et al.": both methods applied at '<60' N or different method applied in 2015 vs 2022? Confusing or something missing?

Line 872: ibicus???

Line 882: Important paragraph. I rewrote it at ~150 words compared to authors' 185 words. See example. ~20% reduction with no loss of info?

Line 889, "maximum VPD as drivers;": I will know VPD as acronym for vapor pressure deficit but many readers will not?

Line 932, Table 5: Header row - Controls rather than Control S? 4th row (header plus 3): "temperature approximated os ISIMIP3a/" - spelling and punctuation errors? 8th row (header plus 7): VCF needs definition? JULES acronym not defined?; 9th row (header plus 8): "tree cover plus none-tree vegetated cover simulated by JULES and bias-corrected as above": capitalize first word (as for other table entries; bias-corrected as above in this Table or in text?

Line 951, "weighted ensemble": 'weighted not explained here? Perhaps later in this paragraph? Authors could rewrite many paragraphs of this section with 15% fewer words.

Line 965, Season Forecast: S2S addresses a very complex, regionally-specific, non-stabile challenge. In any case, at much coarser spatial (and temporal?) resolution than authors need for fire prediction? Authors' figure S1 (in supplement) shows no (zero) predictability for any relevant region of North or South America and only, perhaps, weak correlations for Turkey near Greece. (What do cross-hatches designate in this figure?) So, why do authors waste time and text space on ephemeral low-resolution 'modes' (ENSO, IOD, etc.)? Emulating S2S so they have 'something' to talk about? Better (and, scientifically, more rigorous) to run their well-prepared attribution experiments to see if anything emerges. If something emerged, not valid to then label it ENSO, IOD, etc., at least not without much more work not relevant here? Getting rid of most discussion of these spurious modes would reduce manuscript by several pages?

Line 984, "ongoing debate regarding the direct influence of the IOD on Australian fires': Indeed! If we shouldn't rely on IOD with eastern Australia, why should we accept any other hypothetical correlation?

Line 995, 996, "there are few regions in the word where it is possible to establish statistically significant teleconnection between burned areas and atmospheric modes." Absolutely!! So why have authors wasted so much text on speculation?

Line 1070, "largest contributor to global mean annual totals": statistically speaking, one shouldn't identify large nor small contributions to a 'mean'?

Line 1074, North America: Perhaps a sentence about air quality impacted by Canadian fires here, consistent with South America below?

Line 1075, "contributing": 'contributed' instead?

Line 1096, "Europe: Low wildfire extent in general," Not a valid introductory phrase?

Line 1119, 1120, "Africa: Low wildfire extent in general with BA 13% below average in the African grassland, savannah, and shrubland biome.": Again, not a valid sentence? Again, grassland savannah shrubland distinctions, this time with African inflection? IPCC made definitions not widely accepted by terrestrial ecologists?

Line 1132: reader encountered identical information a few lines earlier. Details here, highlights there? We don't need both? Same comments as above apply here. "MODIS BA product": which, what version, etc. Authors need to settle on clear convenient naming system for MODIS burn products?

Lines 1143, 1144, "African grassland, savannah and shrubland biome, which is the largest contributor to global mean annual BA totals": I appreciate that authors know this but nothing in Figs 1, 2 or 3 proves this point?

Line 1146, "lower BA in savannah-like systems in 2023-24 was not observed in Australia": not clear to this reader whether this constitutes a real decline or a labeling discrepancy?

Line 1175, "58% towards total global BA and 40% towards total global fire C emissions.": this statement needs a certifying citation?

Line 1180, "BA extent was in the top three years on record": because Fig 1 shows only a single fire year, hard to credit this observation from Fig 1?

Line 1205, "prominent regional feature": refers to northern South America but not evident in biome data and not particularly notable compared to Africa or much of Asia. Authors have, no doubt, valid reason for this claim but not evident in their maps? Much of the information presented by the remainder of this paragraph repeats what reader encounters later, e.g. in Section 3.1.3.6 (starting at line 1845). We don't need two repeated versions of same data?

Line 1240, "three good rainfall years have resulted in grass fuel accumulation": here multiple years of abundant rainfall increases fire risk and occurrence? Other places in this manuscript, multiple years of higher-than-average rainfall decreases fire risk? Do authors want to or need to address this discrepancy? If, for attribution

studies, they accept months of wetness (dryness) as indicative of low likelihood (high likelihood) of fire, does this statement expose a weakness in their assumptions? Repeated at lines 1340 & following?

Line 1281 & following paragraph: hasn't reader already received this info? Better here, where more relevant? Not in both places, please.

Line 1282, "moderate resolution satellite data": another term for MODIS burn data?

Line 1297, "higher resolution and higher overpass frequency": but reader learned a few sentences earlier that high-res satellite information confirmed what GFA implied? High-res in this case refers to Sentinel-2 at 10m, compared to VIIRS at ~400m and MODIS at 500m? Words such as 'moderate' and 'high' not particularly helpful in this case. Fig S3 very helpful!

Line 1310, "of the fire also showed high": 'also'? But, reader learned in previous sentence about ~10% difference among various products?

Line 1329: What do authors want readers to learn from this section? That GFA, when it works and when validated by other products, proves reliable? In other cases, however, particularly in urban areas, GFA often fails? Not clear to this reader? When lucky, GFA works great? When unlucky, GFA fails? Plus it detects only larger fires? So?

Line 1335, Review by Continent: Unfortunately, this section reads like WMO annual weather report: warm and wet here, dry there, only station reports, no correlations. Both here and there authors invoke ENSO when they have nothing else to pin to? This reader begins to question value of this entire section. Next year will prove different in detail but no better in overall content or conclusions? Fires always occur. Damage always ensues. Evacuations, suppressions, death: likewise. Nothing new here? "a fairly typical fire season" (line 1406), some places hot, others cold, seems the only useful summary? ESSD should not publish 'fairly typical' reports! But, without this (and WMO reports on weather), where would one find such records? Authors will have saved readers lots of work? Nothing about livestock, please! For this manuscript, readers also need zero info about fire fighting experiences or equipment. For me, excessive detail about Evros fire; how will this info prove relevant to a) other fires, and to b) larger questions about changing fire regimes? Evros fire notable only by BA? So? A fire killed two civilians? Condolences, but what use to a reader to know that info?

Line 1729, Figure 7: Potentially informative but provided here at very poor resolution?

Line 1738, "across the Country that were well in excess" why 'County' not Canada

or country?

Line 1758, "if the USA had also experienced a high fire year, the air quality would have worsened in many states": is this a guess or a statistically-valid prediction as one outcome of this work?

Line 1941: Again, what do authors want readers to have learned from this section? Admiration that authors have gathered lots of info in one place? Granted. Can a single reader assimilate all this info? No. Does this accounting, by continent or cumulatively, make any difference? For this reader: no.

Line 1983, "declines in deforestation rates and deforestation-related fires have fallen": 'declines' have 'fallen'? Silva Junior et al. paper very confusing but title probably best conveys their message "Brazilian Amazon deforestation rate in 2020 is the greatest of the decade". Not clear what the authors intend here?

Lines 2125, 2126, "Some anomalous BA starting in August and extending to November": not a complete sentence, readers need to guess at authors' intent.

Line 2136, Figure 10: Despite good intents and good efforts by authors, this figure not really readable enough to prove useful? $\pm 10\%$ for Canada, $\pm 2\%$ for Greece and western Amazon. This reader does not believe, based on information gained earlier from this paper, that authors can expect readers to distinguish $+1\%$ from -1% . Reader needs some basis to minimize uncertainties here! Resolution of figures in my version = very poor?

Line 2153, 2154, "inadequacies in predicting certain ignition sources or accurately representing fire propagation across vast landscapes in current forecasting systems" A very important conclusion, noted here in case it does not re-appear later.

Line 2161, 2162: "inadequacies in predicting certain ignition sources or accurately representing fire propagation across vast landscapes in current forecasting systems": Another strong important conclusion, conveyed in highlights but repeated in detailed section? Same text!!!

Line 2165, "linked to the strong El Niño.": for reasons already highlighted, including by authors, this reader doubts that authors can assert this link?

Line 2173, "no single factor can explain the most severe fires": another strong positive statement, in highlights, waiting for details to follow?

Line 2199, Figure 11: Very important helpful figure! This reader wonders how a different fire prediction model, something other than Canadian FWI, might have

worked? At longer time scales (longer than 2 weeks or so), no fire model would show significant skill? Authors have skillfully defended choice of FWI, and it seems to make sense given importance of Canadian fires, but they must know about other fire prediction systems? (Thinking here of earlier [perhaps now dated] work with GEE by Gray et al. in ESSD: <https://doi.org/10.5194/essd-10-1715-2018>). Some discussion or assessment? Who else but these authors would know?

Line 2260, Figure 12: Again, informative and helpful. Here, however, reader confronts (for the first time?) 'dead' fuel vs 'live' fuel, both in reference to fuel moisture. If this dead vs live difference proves important, should authors highlight these factors for readers? 'Deap' should instead read as 'deep' in all cases? Lightning basically constant in later months in Amazonia?

Lines 2273, 2274, "lightning as a key source of ignitions in the region": referring to Canada but flash rates much higher in Greece and higher yet in Amazonia?

Line 2281, "vast, densely vegetated": referring to Canada, have authors given readers any bases for accepting this statement?

Line 2293, "intrinsic difficulties in forecasting isolated extreme events": Indeed! Major primary conclusion, but not emphasized by authors?

Line 2329, "anomalous weather conditions subsided later in the fire season": not true, at least as shown in Canadian section of Fig 14? There, weather controls look nearly identical Apr vs Sep?

Line 2356, 2357, "direct human-induced landscape changes exerted minimal influence on the extent of burned areas in Greece": to my eye, human influence (row 5 in Fig 14) looks vastly different Canada to Greece, and very important for Greece?

Line 2361, legend to Fig 14, "increases/decreased": should read as 'increases/decreases'? BA range data, which varies for each site, not clearly evident? Hidden by overlapping graphics? Time spans (X axis) also vary in this figure. Gives this reader even less confidence that I understand the figure? I would have said: a) minor fuel influence at all three sites but different signal Canada to other two; b) strong drought signal in Amazonia; c) solid weather signal at all three sites but perhaps dropping off late in season at Amazonia; d) minor human influence Canada and Amazonia but strong in Greece case? Personally, I don't understand how authors got probabilities to tenths of percent and I find no basis to trust $\pm 5\%$? A highly uncertain model (ConFire) run multiple times on multiple cases does not improve reliability? If Fig 14 presents site-specific BA anomalies relative to site-specific means, but all with same magnitude (same Y-axis units) then I get even more confused? It seems key to me to understand Fig 14 but I don't?

Line 2366, legend to Fig 14, "as a fraction of land area": I appreciate effort by authors to quote and illustrate uncertainties but this reader still does not understand the bottom graphs? Fraction of land area?

Line 2420, Figure 15. Not at all understandable by this reader. The figure legend seems wrong? E.g. this reader sees, for July and June (why this order?), top to bottom: BA anomaly (top), increase from controls (mid), decrease from controls (bot). The legend instead implies left right distributions: "left of the other two maps looks at anomalies in controls that would cause higher BA" while "right map shows drivers that would have led to lower than normal levels of burning". Not correct? June had greater -7 to +7 BA anomalies than July (-3 to +3) so, strictly speaking, user should not attempt monthly intercomparisons? Greyed-out areas should inversely correspond to top BA anomalies, for mid plots, but directly correspond for lower plots? This seems very hard to detect. Then, four size and color dependent dots in each 0.5 degree grid box? I think I understand goals for this figure, but fails for me.

Same problems extend to Figure S13, but for two additional months (Sep, May). Also to Fig S14 for Amazonia. Again: unreadable. Sorry, these old eyes can make no sense of this arrangement. I give up entirely; skip this section.

Line 2492, Attribution. Similar to above: highlights followed by details. At first I liked this arrangement. Now I see it mostly as redundant; largely a waste of space?

Line 2501, 2502, "All forcings combined have led to an overall reduction in today's average BA across Canada." Interesting. But this contradicts all previous assessments for Canada?

Line 2509, 2510, "Climate change has increased today's average BA in the Mediterranean region, but this has been mainly offset by socio-economic factors.": reference Greece but same conclusion as for Canada: socio-economic offsets climate leading to no net change?

Line 2517, 2518, "Climate change has increased today's average BA in the Northwest region, and all forcings have led to an overall increase in burning". In this conclusion, western Amazon differs from Mediterranean and from North America?

Line 2529 to 2531, "probability of experiencing the high fire weather observed during June 2023 is more likely in a climate forced with anthropogenic emissions.": Referencing high BA across Canada, this conclusion (plus Fig 17) contradicts what reader found in highlights?

Line 2578, 2579, "BA in Canada in June 2023 was 0.8-38.0% greater due to total climate forcing in the 2003-2019 period" Perhaps labels have confused this reader, but if 'total climate forcing' = 'All', then this conclusion contradicts what we read in highlights?

Line 2599 to 2601, "uncertainties around whether total climate forcing and socioeconomic factors caused an increase or decrease in BA are higher, and the smaller region size makes detecting a strong signal of change more challenging". About Greece, safest to draw no firm conclusions?

Line 2610, 2611, "additional burning could have been up to 0.8-36.2%" but, from Line 2615 "socioeconomic conditions (95.98% likelihood) increasing BA by 0.18-5.32%. Authors lost this reader long ago on how to trust 1 to 36 vs 0 to 5. In any case, for western Amazon region we can only say with confidence: perhaps?

Line 2623, legend for Fig 18. No explanation/assignment of orange color? Oops, now I get it: orange = factual, blue = counterfactual, combination (overlap) = purple. Not immediately clear? Taking western Amazon, this reader finds 1.53 to 7.66 (2 to 8) modest increases in high BA (how authors derive 99% confidence for curves that basically overlap remains beyond me) vs 0.18 to 5.32 (0 to 5) increase in socio-economic factors. These numbers DO NOT accord with those at lines 2610 and 2615 in text. This reader now concludes, using corrected numbers, no net change for western Amazon in high BA?

Line 2667, legend to Figure 19: In this Fig, authors did not use transparent colors so colors do not blend as they did in Fig 18? If authors did not present Y-axis (probability) in log scale, users would see nothing? I figured out what bothers me about these ultra-precise uncertainty estimates: authors compare final products of long train of manipulations, each with its own (sometimes acknowledged, sometimes not) uncertainties, source observations (satellite or reanalysis), aggregations or disaggregations, final manipulation for effective presentation, etc. but only report (valid?) uncertainties deriving from the last step.

For ESSD, we need full end-to-end, uncertainties! A full uncertainty budget! For satellite data: orbit changes, sensor degradation, processing levels, cloud or aerosol or competing species absorption, etc. For reanalysis: data ingestions, processing, discarding, spatial inhomogeneity, etc.; reanalyses derive from need to produce error matrix! Processing: use of anomalies, use of relative anomalies, changes in axis ranges, etc. Uncertainties at every step, propagating, canceling, amplifying, whatever. To report only the statistical inter-comparison of final steps (e.g factual vs counterfactual) ignores a host of uncertainties inherent in how authors got that far? I would rather not check all numbers (although prior comments relative to Figure 18 suggest that someone needs to validate), but I

react negatively, as will most readers, to tenths of a percent in final precision when readers know and authors well know that basic uncertainty proves higher in every case! How can authors report -0.2% differences between model outcomes when we and they know that we can't even measure BA from satellite to better than $\pm 10\%$? I would feel very pleased to get proven wrong on this, but, with no disrespect intended for authors, they make much in these sections of very tiny differences only by ignoring much larger underlying uncertainty! ESSD readers need to see underlying systematic end-to-end uncertainty budget. Very difficult to produce, especially in this case? Perhaps, but tell us. Might mean that one can not, in fact, distinguish factors that influence max BA or median BA? Tough news, but tell us; we need to know. And we need to hear recommended solutions from these experts!

Summary: great respect for efforts and reports, but manuscript lacks an encompassing uncertainty budget.

Line 2674, Seasonal and decadal: Again, highlight followed by details, leads unfortunately to substantial redundancy.

Line 2678 & following paragraph, IOD and ENSO: But, authors stated earlier that no statistically valid links exist between BA and atmosphere modes (line 995, 996). Why then do readers confront more discussion of IOD, ENSO ("ENSO-neutral"), etc?

Line 2688 & following paragraph: But readers learned in previous paragraph that (for Canada) "no clear signal for extreme anomalies is present". Why then does this paragraph focus on scenario differences (SSP585 vs SSP126) and mitigation actions. To remedy signal that does not exist?

Line 2701, "above PI": readers will not understand PI without definition?

Line 2704, 2705, 2707, 2708: as if readers did not already suspect "projections show a high level of uncertainty in all regions" and "we cannot determine if current mitigation efforts are effective". Valid, honest conclusions, but don't those render this entire section moot?

Line 2745, legend to Fig 20. If authors already concluded lack of statistical connections, and one of our best modeling groups predicts this wide (useless?, >3C in every case?) range of near-future temperature ranges (one for ocean SST in central equatorial Pacific region, the other we would have to search for) why should reader give any credence to any of this? Prove challenges by showing current weak model skill but don't waste our time with speculations?

Line 2679, legend to Fig 21: Based on careful reading of all prior guidance, this

reader has learned to disregard this information entirely. July & August for 95th % basically empty. Reader could easily have guessed at areas for 75th in July and August? Nothing against ECMWF, just that we should have learned by this point in this manuscript to trust no intense FWI probability < 100%, which basically don't occur anywhere on the planet after June?

Line 2780, future changes section: Direct mapping of BA pixels to model output pixels? No intervening fire model, predictive or attributional? Methods at Lines 1030? JULES never defined? But figure time extents do not accord with method-specified boundaries. Fig 20 definitely not, Fig 21 perhaps but only for Canada?

Line 2786, 2787, "ISIMIP3a and ISIMIP3b": this reader understands a few differences between 3a and 3b but authors provide this and most readers no help and no guidance?

Lines 2796, 2797, "likelihood of a 2023 fire event recurrence increases to 0.31% - 0.9% - two to almost four times as likely as today (Table 7; Figure 22). But, remains highly unlikely? This outcome (these numbers) very hard to extract from Table 7 - one needs to confuse min and max or cross scenarios - so data come from Fig 22? This reader unfortunately gets nothing from Fig 22, particular at level of detail implied here? In next 10 years? Nothing.

Line 2798, "different SSP scenarios diverge significantly by 2070": Really? By what criteria? Lack of range overlap, as authors specified at line 2790? Very hard to confirm authors' conclusions from Fig 22!

Lines 2802, 2803, "at least one event similar to or worse than the 2023 event occurring again is estimated to be between 16% and 25%": If reader hopes to confirm these author-specified changes (at least these seem more moderate), where should they turn. For Canada in Fig 22, likelihood range only goes 1 to 5%? Does one need to convert and guess from BA event numbers? Not acceptable.

Line 2804, legend to Table 7: No explanation of colors? Obviously related to magnitude, but how? Many cells report insignificant results (48 of 160, 30%)? Predominantly not valid for W Amazon, regardless of return rate (1 in 6 vs 1 in 100)? Marginally valid for Greece? Please fix wrapping and display for 2nd column. Entire table needs improvement! Minimum likelihood outcomes for Canada 'increase' from 0.08 to 0.31? Readers should learn what from this? Finally, here, a definition of 'extreme', in a UNEP report? But, in text later (lines 2845, 2846), UNEP 2024 seems to specify something different? This reader does not want to search for UNEP 2024; we need authors' best explanations!

Line 2825, "changes in both [djc: fuel load and fuel moisture] controls are needed to explain divergence between SSPs": very hard, perhaps impossible, for reader to

credit this conclusion based on these data!

Line 2828, legend to Fig 22: No indication of when or if signals emerge from uncertainty bounds, either in time or in event return times or % likelihood? No explanation of means, medians, whatever, indicated by lines within each bar? This reader guesses: perhaps a scenario difference by 2100 in Canada, driven perhaps by fuel abundance? Before that, nothing significant? For Greece, perhaps significant differences by 2090 and 2100, driven by moisture? For W Amazon, nothing (zero significance) for any decade or any scenario? But these personal conclusions diverge substantially from authors' narrative? Here they compare ConFire simulations to re-analyses? No indication of large uncertainties in each?

Line 2823, legend to Fig 23: Not the US-Canada border that I know? Even if outlines stay north of Great Lakes (border actually intersects lakes via mostly straight midpoint lines), remainder to west looks step-wise distorted whereas actual border remains dead (latitudinally) straight, Minnesota to Vancouver Is? Relevant because only signal seems to occur along southern border, near N Rockies and/or Glacier NP. Surprised to see, in col 1, an apparent minimum along southern border? (Do authors think readers will have any concept or trust of 0.001% difference in BA? Or any concept of 0.05 vs 0.5 % change?) In middle and right columns, at least for SSP585, fire extents and likelihood of return seem to max in that area? This reader develops no confidence in any part of this figure? Dot and change scales not linear? Arbitrary?

Line 2844, "some areas" ... see increases ... "almost everywhere": No sense in that statement. Authors expect users to accept that areas of low BA today will become areas of preferred BA in the future? Authors or ConFire or both haven't the faintest idea?

Lines 2862, "from 1.3% to 0.67-1.78%": relevant to Greece, authors expect users to accept or understand significance of such a negligible change? Later (line 2872) authors report "no significant difference between the scenarios by 2100. Why, then, have they wasted user time and their own page budget?

Line 2889, legend for Fig 24: Same questions as for Fig 23, but why? Speculation? Waste of valuable space?

Line 2941, Summary: Useful organization here: paragraph responses to each original objective. But these paragraphs read mostly as redundant restatements of previously-discussed results? Reader does not need both? This section seems better organized and more concise?

Line 2934, "suggests": authors can do no better than 'suggest'? Reader conclusion: nothing significant, why waste my time?

Line 2993: here, reader confronts mention of "Alexandroupolis" fire, A few lines earlier, authors mentioned "Evros" fire (line 2972). Two names for same event?

Line 3012, "release break": New term for readers? Not used previously?

Line 3044, "socioeconomic factors outweigh the increase from climate change", many readers will agree with this statement, even sans hard evidence, but preceding paragraph seems to express contradictory opinions: climate change increases probability of large BA while socioeconomic inhibit? Decide what you want reader to take from this paragraph?

Lines 3055 to 3057, "At extended lead times (greater than 2 months), ... no discernible signal about moderate or anomalous conditions is identified" Very worthy and valid conclusion, but negates much of what authors worked to present? Earlier in this paragraph, more vacillating speculation about ENSO or IOD, etc?

Lines 3060 & following paragraph: Cautious but confusing? From final sentences, this reader concludes: even with best current mitigation options, warmth coupled with drought (interspersed with floods) will increase fire vulnerability? I truly worry that better weather, soil, vegetation data in Canada influence all these Incomparisons!

Line 3076, Frontiers: I and most readers will welcome and agree with these concerns. Definitions, observations, predictions, attributions, future projections. Too long in every case but see recommendations above. Very valid helpful cautious summary, ideal coming from this group? Mention SI? Also a place to encourage other contributions to proposed or other future special issues?

Line 3393: Frontiers. Good section! Here, particularly, you need to focus on research audience or fire management audience? Way too long! Could reduce by 30% with no loss of information (see above),

Line 3430, "extensive tree mortality": First (and sole?) hint at infestations. For much of northern Rockies, as for many forests in Germany, climate-enhanced (activity, reproduction) insect infestations kill 1000's of trees. At very least, this large-scale die-off moves vegetation from 'live' to 'dead' categories, a distinction invoked once or twice but not much discussed here? Impact seems to involve climate, weather, vegetation and human controls? Where would this fit in existing modeling? Or in future needs? Insect-mediated factors could grow?

Line 3460, data about fires and fire impacts: in an earlier ESSD paper, Karen Short of Montana Forest Service reported difficulties in rectifying versions of fire data

into one accurate valid record. As I remember, nearly 40% of fire reports in her US-based database proved invalid or redundant. Emphasizes dependence on remote sensing data? Strength or weakness? Also emphasizes that we currently cannot, at least not in near-real time, furnish local experts with accurate local data? Causes us to rely on media, which have even worse unstable biases? Also raises issues of industrial (including transport) fires. Not currently monitored or logged? Often, in media accounts, prove sensational. Notable impacts on emissions, air quality, evacuations, human safety, etc? Another aspect of fire currently out-of-bounds for fire researchers?

Line 3490, air quality impacts: ground-based measurements, dispersion and transport studies, plume dynamics, animal-based human exposure/response models? Add intermittent ill-defined fire emissions? Almost an impossible problem?

Line 3526, authors here tend to focus on fire management, by federal or state agencies vs IP&LC communities. Good points. But, Canadian projections (e.g. from Fig 23) pass through - on both sides of Canadian-US border) - large reservations. Smoke air quality exposures might prove adverse in those communities and locations?

Line 3564 & following paragraphs, economic impacts: omission of or ignorance of industrial fires might prove key here? Economic impact estimates miss many long-term health costs already mentioned?

Line 3740 & following paragraphs: Back to observations, predictions, attributions, future projections. Good stuff but serious redundancies with Frontiers Section. Too long but authors need to think here about readers: what new recommendations do you want readers to remember from this section that we did not already learn from Frontiers section?

Lines 4982, 4983: URLs (e.g. <https://doi.org/10.xxxx/longer> number that wrap across lines don't work as hot links.

Perhaps now

Need a table of contents? Or, something different or better to help readers find specific information. Perhaps accompanied by short paragraph explaining what you keep in main manuscript and what additional info users can find in supplement. Supplement would also need a table of contents or equivalent?

Document would benefit from list of acronyms? Developing such a list might help authors check & track such a long list of acronyms?

Line 102: "extreme" fires? As authors repeatedly point out, community presently confronts inability to reliably quantify nor communicate 'extreme? Should these authors apply a term they later, on valid grounds, question?

Lines 131, 132: "insights relevant to policymakers, disaster management services, firefighting agencies, and land managers," But, this important target group of readers will find 124 pages daunting or forbidding?

Line 133: Recommendations - well-grounded - that follow cover basic understanding as well as "preparedness, mitigation, and adaptation"?

Line 258 & paragraph following: initial caution & questions about what constitutes extreme events arises here. Good! But, authors have already used terms 'extreme', 'extreme fire', 'extreme wildfire', leaving this reader with two questions. A) Should one use such a complex fraught term? B) What will these authors, best in their field, suggest as alternate or as quantitative?

Line 268 & following "Objectives": Good section, appreciated. Shorten? Opening paragraph suggests tools and information (good), prospective, but user just read that authors will focus in BA while first objective definitely leans toward recent observations. Somehow, perhaps in introductory paragraph of this section, remind users about intent to apply and report data and observations?

Line 326, paragraph starting here: Good paragraph, first discussion of uncertainty in application of term 'extreme', very useful caution, but authors could clean this up by 10% to 15%?

Line 396 & paragraph following: Important paragraph outlining timing assumptions and choices. This paragraph should perhaps move to top of Methods section?

Line 701 & following, Section 2.3.2.3 on Drivers: authors could rewrite this entire section with 15% fewer words but no loss of info?

For future discussions and planning

Move everything not global nor related to specific focal events to supplement?

Authors have, in this version, 'teased' readers with three specially-chosen focus events. Because each of those focus events includes long discussions of present predictability and future probability, won't readers want, in second version, some update on authors' conclusions for these regions? Meanwhile, authors will select a different three regions for 2024-25? Eventually, doesn't one end up with oldest, older, prior and present events?

Background, global, regional, predictions: different tools and different conclusions for each. Should these evolve into separate products? Mention here ideas about a special issue? Perhaps parse under several journals, according to?