Peer Review File

Manuscript Title: Biodiversity impacts of the 2019-20 Australian megafires

Reviewer Comments & Author Rebuttals

Reviewer Reports on the Initial Version:

Referees' comments:

Referee #1 (Remarks to the Author):

A. This is a study where the authors collected a huge amount of data on species occurrence and abundance after the 2019-2020 Australian fire season which was the most active fire season in recent memory. They found some generalizable patterns of resistance to fire among many taxa, with implications for fire and vegetation management.

B. The dataset has a broad geographic and taxonomic extent, making it fairly unique, especially given the post-fire nature of it, and that many datasets also had pre-fire data. So the scope of the field-collected data makes it of interest to most ecologists.

C/D. They used a meta-analysis approach. I saw a minor flaw in the statistical methods that would be easy to fix and would not change the results substantially (model selection with ML instead of REML). The one thing that I don't really understand is how they did the splitting of positive and negative effects. That needs to be explained better to make an informed judgement of the paper. Another thing about the approach that puzzled me is that there were no open ecosystems like grasslands and savannas considered in the analysis.

E. The conclusions are generally robust and valid, but the confusion I had on the aforementioned splitting of the effects
prevents me from really staking a claim.

F. Line by line comments below for suggested improvements

G. References were appropriate

H. I thought it was generally well written, except for a few areas mentioned in the line by line comments.

---line by line notes:

recommendations of increased fire suppression a good countermeasure to the western US based prevailing narrative

154: Wildfire burned area is actually decreasing globally (Andela et al 2017; chen et al 2023) due to urbanization and agricultural intensification in Africa

190: occurrence OR abundance? which one when? Maybe add "(See Methods, section X)"

197: (also 819-821) This needs a little more explanation. I find myself struggling to understand the meaning of the blue and black error bars, since some of the blue ones are negative and some black ones are positive. I don't get it.

201: "biodiversity" not defined, is this species richness? important for the interpretation of the effects

215-216: This is pretty vague

222-224: Were these frequently burned areas places that are adapted to frequent surface fire, and then burned at high severity (i.e. stand-replacing)? Might there be a reason to believe that places that hadn't burned were adapted to infrequent, stand-replacing fires? Low severity had positive or neutral effects on places with 3+ fires since '79, suggesting that areas with a high-frequency fire regime were less impacted when severity was low.
High severity had neutral effects but overall negative on places with 0 fires since '79

355 -65: probably can't draw too many conclusions about plants, since an assessment of the recovery of trees, shrubs and perennials might necessarily require 5 or more years after disturbance

Fig. 4: It seems odd to me that there are no grasslands? It seems like it might be good to include another landcover-based effect size estimation, with open-canopy ecosystems like grasslands and savannas versus closed-canopy systems.

438: this study really is looking at resistance to fire, and resilience to fire is probably not able to be assessed from this data.

466: I might add joint species distribution models

738: Typo: "unburntburnt"

793-4: "test = "t"; dfs="contain"; method = "REML" 69. We also set control=list(optimizer="optim", optmethod="Nelder-Mead")". write out what all this means instead of copy and pasting R code. e.g. we used a T test and restricted maximum likelihood, with the Nelder-Mean method of model optimization. Make a github repository that fully reproduces the analysis and refer to that for folks that want to see the code.

808-813: it looks like you're doing some model selection. In that case, I think the proper way to do it is to use maximum likelihood to compare models, then use REML for the chosen model. This probably wouldn't change much, but it seems like it might be a good idea to make sure you do it like that.

834: Still do a github repository that has all the code from the analysis along with links publicly available data used, so that someone making reasonable request for the data can repeat the analysis.

general methods comment: might be good to italicize r package and function names.

Referee #2 (Remarks to the Author):

This ms presents the results from a meta-analysis of impact studies related to the massive 2019-2020 fires in Australia. While it is based on a unique dataset and reveals potentially important findings, I found the presentation of the work to be problematic and less compelling than it could have been (IMO). I hope the following points are useful in crafting an improved revision.

1) Abstract (and throughout): There is a general emphasis on the negative effects of fire, in the interpretation of findings, despite often reporting strong positive effects of fire. This is a bit confusing and at times feels like “cherry picking” results to support specific management and policy positions.

2) Line 201: Beginning here the authors essentially refer to “biodiversity” as the dependent variable, when it is actually not species richness (or some other common metric of biodiversity) that is being examined. Instead, it is really reflecting occurrence or abundance of whatever plant or animal species happened to be part of the studies in the meta-analysis. I’m not sure what better term would fit here, but we are not dealing with “biodiversity” as it is usually used in scientific research and conversation.

3) Line 212: The first sentence here is quite circular – “More severe fire is expected to have more severe ecological effects” – as “severity” is an ecological concept. It glosses over what characteristics of a given fire lead to more ecological impacts. For example, sometimes frequent fires are ecologically severe, while in other systems it's high intensity fires that have severe ecological impacts. In still other systems, it may be larger fires that are most severe. What fire characteristics are being examined? If we're talking about heat release or canopy consumption, it would be helpful to be more specific, without going deeply into the methods here.

4) Line 219-220: The first sentence about short interval fires is generally not true for grasslands, which can reburn very often, nor for many of the world's mixed conifer forests that (prehistorically) burned often at low severity (i.e., consuming understory and benefiting dominant overstory trees). HOWEVER, the notion of "fire deficit" in western US forests is so widespread that many will find this sentence to be problematic. More nuance/context needed for international readership.

5) Line 228-230: Again, although potentially important and true in this context, it is in opposition to the

"fire deficit" bandwagon mentioned earlier. More nuance/context needed for wide audience.

6) Line 232-240: This paragraph reports positive effects – seeming to contradict the previous paragraph – without the context needed to understand why they occurred. Further, the reporting of both positive and negative effects, for largely the same variables, happens repeatedly throughout the rest of the ms. You offer some literature and/or hypotheses as to why different effects might be observed, but this whole presentation is not very satisfying without a deeper understanding of WHERE the effects are being observed. Figure 2 is also increasingly frustrating and unsatisfying as a key way to show us what is going on.

7) Lines 243-245: The authors assert that "Conservation and management in fire-prone regions must therefore focus on maintaining areas with intermediate inter-fire intervals, where we observed the smallest effect sizes (Fig. 2c)…" This seems counter-intuitive. Why wouldn't you instead aim for the fire intervals (and severities) that coincided with the biggest positive effects on biodiversity, in those places where that was relevant. In contrast, in environments where a certain type of fire had big negative effects, you would avoid using that kind of fire regime. Why aim for the smallest effect sizes? It seems like one of the important insights of this study is to identify which types of fire regimes are best for biodiversity in which parts of the study area... Again, the spatial aspect is missing.

8) The effects figures quite difficult to read and absorb, including the interpretation of P values. Are all of the "Reduced" and "Increased" significantly different from zero, while only a few of the "Overall" are significant? What does it mean when the "Severity P" is not significant in some panels?

The text does not seem to differentiate by significance when reporting results... why?

Are the "Overall" effects ever really meaningful, and are they worth reporting? It's where the positive (or negative) effects occur and why that is interesting, no?

9) Lines 289-299: Again, it is confusing and causes a bit of "whiplash" as one finishes a relatively compelling paragraph on negative effects (on plants? animals?), but then is presented by largely contradictory results in this new paragraph.

10) Lines 294-299: While some of these hypotheses are interesting, we are looking at a very recent fire and only ~3 yr of recovery. It seems very unlikely that these results are reflective of longer term recovery patterns, making their relevance questionable for wider management and conservation questions.

11) Lines 302-305: This seems like a prime opportunity to show WHERE the positive vs negative effects were observed, no? Maybe I'm not understanding something about the analyses, but the sites are shown in Fig 1, so mapping out the results would help greatly with interpretation and could also reveal additional variables that could be included in the analysis.

12) Line 305: The “buffering effect” of protected areas highlighted here and later on is an unusual interpretation. Granted, effects are generally less in protected areas, but they are also not experiencing the 26% positive effect in some areas either. Is this a good thing? Where the positive effects are happening (and to which taxa) is important information, and knowing this would impact the entire ms.

13) Lines 248-253: How do these results compare with other longer-term studies across animal taxa? For example: https://onlinelibrary.wiley.com/doi/full/10.1111/ele.14177

14) Lines 400-411: You emphasize the negative effects here, but at nearly every turn you present substantial contrasting negative AND positive effects.

Is there a reason to discard most of the positive effects and focus strongly just on the negative? Could you make the case that the positive effects are fleeting or artifacts and thus less useful? Can't you use the location of the sites to better inform a mapped interpretation? (Sorry, broken record…)

15) Lines 432-435: Where are these refuges within the study area, and how did they fare?

16) Table 1: This table is useful for the reader, but it feels a little late in the game to make sense of all of the preceding (apparently) contradictory results. It could help to give a shorter overview of such mechanisms earlier in the paper.

Why isn't a table like this based on the large literature that exists on how fire interacts (positively and negatively) with plants and animals?

Bond & Keeley, 2005; Carbone et al., 2019; He et al., 2019; Keeley et al., 2011; McLauchlan et al., 2020; Moritz et al, 2023; Nimmo et al., 2019; Pausas & Parr, 2018; Santos et al, 2023

17) Line 457 onward: The proposed mechanisms could be supported by, or even generated by, mapping out where you found evidence of such pathways, based on the sites in the study and the effects observed. This would be a much more compelling way to justify the mechanisms presented in this table. I would suggest adding such an analysis, even if relatively superficial.

Referee #3 (Remarks to the Author):

My review is focused on the dataset and the meta-analysis.

Dataset

- The authors performed a meta-analysis in order to estimate impact of mega-fires on biodiversity. However, their meta-analysis does not seem to be based on a systematic review of the available studies but to a set of "projets" whose origin is a bit unclear. Although the list of projets is provided, the authors did not explain how these projects were retrieved. The authors should have implemented a systematic search strategy and described it in details in order to ensure transparency and reproducibility. They should explain how their search strategy was designed to ensure that all relevant projects were retrieved.

- In addition, I was wondering why the authors focused in Australia only, while mega-fires occured in other areas as well (e.g. California). Because of this choice, it is difficult to understand how their results could be extrapolated.

- The authors listed a series of selection criteria (lines 710-719) but did not justify them very clearly. Moreover, they did not explain how the data were extracted.

- It would be useful to produce a diagram similar to the PRISMA diagram (a standard diagram in systematic review and meta-analysis) in order to present the results of their review and selection procedure.

- More generally, I advise the authors to consult standard guidelines on systematic review.

Effect size and meta-analysis

- The authors used different types of biodiversity measures and they mixed occurence and abundance of taxa together. It would be interesting to check whether the results are different between these two types of metrics.

- 723: do you mean « one to five » ?

- I don't fully understand why the authors used standardized mean difference (SMD) and not (log) ratio as an effect size. The problem with SMD is that the differences between groups may be due to either difference between mean values or difference between standard deviations. This makes the interpretation of the results a bit difficult.

- In before-after studies, the authors used the standard deviation « before » and not a pooled estimate of SD. Why?

- How the differences in sample sizes were taken into account? Maybe all samples had the same sizes, but this is not fully clear.

- The standard deviations were computed using different methods, depending on the protocols of the projects. It is unclear why another method was used for control-impact studies. Why not using the same method for all studies?

- The authors excluded the zero data (no difference between burnt and unburnt). This can lead to a biased estimation (overestimation of the effect). The authors mentioned convergence problems, but this is not a good reason for removing data. In case of numerical problems, the authors could either add a small constant to the data or used a different statistical model able to deal with zero data (or use both strategies and compare the results). At least, the authors should estimate the probabilities of zero with a separate statistical model (that may even include covariates, e.g., glmm). Then, the authors could either present the results of the estimated probabilities separately, or combine these estimated probabilities of zero with the mean effect size estimated from the non zero SMD data in order to derive a global average.

- Concerning the control-impact studies, it is not explained how the burnt and unburnt sites were paired together.

- I didn't understand how the SD values were computed for the third type of design (756-748).

- 762-763: as one single control was paired with two different treatments, the two resulting effect sizes were not independent. However, the authors took this non-independence into account. This is a good point.

- 826-832: Here, I understand that the authors used another selection criteria (risk of confounding with drought). This should be explained before, where the authors described the selection procedure. Moreover, it is unclear what the authors mean when they write « we found some evidence ». The authors should use a transparent rule for study selection.

- It would be useful to check the influence of the data with risk of confounding on the results and check the robustness of the conclusions. I would thus redo the analysis with this group data and compare the results with those obtained without. Alternatively, you can add a covariate related to the risk of drought in your statistical model in order to deal with this issue.

Referee #4 (Remarks to the Author):

Overall, this paper represents an impressively large-scale and somewhat standardized collection of ecological data examining the effects of the 2019-2020 Australian fires on biodiversity. The information contained in this paper has the potential to be uniquely influential scientific knowledge that will be a critically important contribution to the literature. The methods used are relatively standard, non-controversial meta-analytic approaches, and seem to be implemented correctly.

A consequence of the ambition of the study is that a staggering breadth of empirical results are included in the paper, which is both useful but also overwhelming to the reader. The paper is challenging to read and follow simply because of the sheer number of comparisons and tests that are examined. As such, the paper appears as scattered and unfocused. Of course, I recognize that distilling the paper any further than it has already been distilled will be a challenge.

The authors idea to split up negative and positive effects was a good once, so as to avoid incorrect inferences of “no effect” when the individual effects are highly variable and include many positive and negative effects. However, a consequence of this decision is that the communication of effects becomes highly convoluted. For example, beginning at line 318: “Mean negative and positive effects were larger at high compared with low severity for Amphibians…” There are so many qualifying words in these types of statements that it is really challenging to follow for a reader.

Perhaps another part of the difficulty in tracking and interpreting the results is they are visualized in three very busy and similar-looking figures, each of which contain many rows of numbers, labels, colors, abbreviations, nested components, p-values, etc etc. I do not envy the challenge the authors face in visualizing the results, which again, are staggeringly extensive. The effect of the presentation of many dozens of tests and similar figures is that the paper is highly repetitive, and lacking in terms of synthetic take-home messages.

I would encourage the editors and authors to work together to try and find a path forward that allows all of this important information to be accessible to readers while providing a more synthetic, focused narrative in the main text that more effectively communicates key results to readers.

Specific comments:
145: “Fire management should aim to defend…” This statement comes across as an advocacy position. Could be rephrased to be more stance-neutral: “Fire management would most effectively conserve biodiversity by safeguarding fire-sensitive ecosystems…” (for example)
157: “…on record, generated a record…” consider rephrasing
249: How was the 2.5 km buffer size selected?
399+: In these section, I would love to see a few more linkages to how implications, fire management recommendations, etc., might be similar or differ for other systems. The effects presented here, while extensive, are limited to one region of Australia. Some or many of the recommendations will not transfer to other systems. It’s important that this is discussed at least briefly.

Author Rebuttals to Initial Comments:

Response to referee's comments on Driscoll et al Megafire impacts are mediated by drought, protected areas, and fire regime

Author's Overview

We have really appreciated the input of the four referees. In responding to the suggestions and queries, we have been able to vastly improve the manuscript. Some of the key improvements include:

- We have more clearly explained why we report negative and positive effect sizes.
- We have added a new conceptual figure, along with a new way of presenting the main results which has made it much easier to understand the results, without oversimplification.
- We have provided a more rigorous justification for our use of standardised mean differences, but also draw on results from a new analysis using mean differences.
- We have more clearly articulated why our work is globally relevant.
- We have made it clearer that we do not use systematic review, but instead use an innovative approach based on collating site-level data across 62 separate projects.
- We have provided additional analyses and summary statistics from repeating our analysis using different combinations of data and different effect size calculations.
- We have added new covariates for (1) region, to examine geographic variation in effects, and (2) response, to examine differences between occurrence and abundance data.

Detailed responses to all referee comments are provided below. In our response we have retained any referee-provided numbering, although some individual comments have been subdivided to better address each point. Our responses are in blue text and quotes from the revised manuscript are in italics.

Referee #1 (Remarks to the Author):

A. This is a study where the authors collected a huge amount of data on species occurrence and abundance after the 2019-2020 Australian fire season which was the most active fire season in recent memory. They found some generalizable patterns of resistance to fire among many taxa, with implications for fire and vegetation management.

B. The dataset has a broad geographic and taxonomic extent, making it fairly unique, especially given the post-fire nature of it, and that many datasets also had pre-fire data. So the scope of the field-collected data makes it of interest to most ecologists.

R1-C/D. They used a meta-analysis approach. I saw a minor flaw in the statistical methods that would be easy to fix and would not change the results substantially (model selection with ML instead of REML).

Author Response:

Please see our detailed response to your comment R1-13 below about using ML and REML, where we explain that ML is not needed because we do not use model selection. Model selection cannot be used because effects are based on different subsets of the data.

The one thing that I don't really understand is how they did the splitting of positive and negative effects. That needs to be explained better to make an informed judgement of the paper.

Author Response:

Ideally in a meta-analysis, the data represent a homogeneous process, and the overall mean of multiple effect sizes converges on a single best estimate of the effect size. In ecology in particular, multiple processes are in play, and some of these can be accounted for using covariates. However, there are always winners and losers after a fire, patterns long represented in post-fire succession research. There are early successional species that increase rapidly after fire, and late successional species that decline abruptly during or shortly after fire. The winners and losers are responding to different processes and so can vary independently. Our main big picture summary of trends, now illustrated in a new Fig. 2, emphasises that larger negative and positive effects occur under more extreme disturbance. The typical overall mean from a meta-analysis would tend to cancel out and therefore mask these effects; the mean of large positive and negative effects will be similar to the mean of small negative and positive effects.

We have edited the last paragraph of the methods section to more clearly explain this:

"Species are expected to have varying responses to wildfire, with some declining and others increasing shortly after fire. Meta-analytic models fitted to plant and animal responses to fire are therefore not modelling a homogeneous process. We applied our meta-analysis to all effects, giving the overall mean effect sizes. However, to better understand the effect sizes of taxa that increased or decreased in response to fire, we used a novel approach which also applied our analysis to just those effects that were less than zero (negative effects), and, separately, to those effects greater than zero (positive effects)."

We believe this approach gives a more nuanced view of the range of effects that fire can have across taxa, and different ecosystem contexts, which is not only interesting from an ecological perspective, but of great concern for policy-makers and on-ground managers. Mean effects, especially related to fire, run the risk of leading to generalisations that ignore major declines of some species and increases in other species, but these generalisations might nevertheless be adopted in management.

Another thing about the approach that puzzled me is that there were no open ecosystems like grasslands and savannas considered in the analysis.

Author Response:

Our study was limited to the ecosystems that burnt in the 2019-20 Australian bushfires, which occurred in woodlands, forests and other embedded ecosystems. Grasslands are rare in the bushfire zone, and even though we combined them with heathlands (Extended Data table 1), there were too few data points to include in the broad ecosystem analysis. We have extended the caption for Extended data table 1 to explain this:

"For the ecosystem covariate, level grass_heathland had five or fewer effect sizes in each analysis, precluding reliable estimates of mean effects, and causing models to fail to converge, so this ecosystem group was excluded when ecosystem was fitted as a fixed effect. Similarly, there were five or fewer wetland sites burnt at high severity, so these were also excluded from analysis when ecosystem was fitted a fixed effect."

Savannas occur in northern Australia, outside of the bushfire zone of interest, which we defined in the methods with reference to Ward et al. [1]. Figure 1 also shows that the study area does not include the northern savanna regions. While there were many fires in the northern savannas, these were not anomalous. To help clarify this, we have added a sentence to the first paragraph of the methods:

"While there were many fires in Australia's northern savannas, these were not anomalous and so are not part of the 2019–20 fire footprint (Fig. 1)."

R1-E. The conclusions are generally robust and valid, but the confusion I had on the aforementioned splitting of the effects prevents me from really staking a claim.

Author Response:

We hope our explanation above has clarified how we used the negative and positive subsets of effect sizes to calculate mean negative and mean positive effects.

R1-F. Line by line comments below for suggested improvements

R1-G. References were appropriate

R1-H. I thought it was generally well written, except for a few areas mentioned in the line by line comments.

Author Response:

We have addressed all of the line-by-line comments, detailed below.

---line by line notes:

R1-1. recommendations of increased fire suppression a good countermeasure to the western US based prevailing narrative

Author Response:

This is a good point. R4-8 also raised a similar wish for more contextualisation. We have added a new second paragraph to the revision in the future management section:

"Increasing use of rapid wildfire suppression is likely to apply to other regions of the world that experience vast fires such as Canada in 2023 [2], and in ecosystems that are threatened by fire, such as rainforest[3]. Some forest types, like Australia's ash forest[4] or North America's coastal temperate forest[5] could benefit from rapid wildfire suppression continued over many decades or centuries, consistent with historic fire regimes. But this will not be the case everywhere. Some dry forests of western United States can be restored by adding frequent, low severity fire to maintain an early successional understorey [5]. Further, many USA forests can benefit from allowing mixed severity wildfires to burn [6]. With increasingly severe fire weather, a range of counter measures are needed, tailored to specific ecosystems. Indigenous fire practices, placing the right fire regimes into the right country [7], is therefore increasingly important for improving fire management for biodiversity conservation [8]. Complementing these ecosystem-specific approaches, we expect rapid wildfire suppression will need to be used more often to limit the extent of land burnt at short intervals."

R1-2. 154: Wildfire burned area is actually decreasing globally (Andela et al 2017; chen et al 2023) due to urbanization and agricultural intensification in Africa

Author Response:

While Andela et al 2017 report global area burnt has declined due to increasingly intense land use, we think it is important to keep the focus on the growing wildfire crisis engulfing many parts of the world. We have therefore changed our statement to refer to 'many parts of the world' rather than 'globally' to avoid conflicting with Andela et al 2017, and added reference to globally worsening fire weather, reported by Ellis et al 2022, and predicted worsening conditions (Senande-Rivera et al 2022).

The sentence now reads: *"Wildfires are increasing in many parts of the world and fire weather is becoming more severe globally due to anthropogenic land-use and climate change [3,9,10]. Recent examples of extreme megafires are reported from the Americas,* *Mediterranean, Russia, and Australia [11,12]. Trends of increased fire risk are predicted to continue, particularly in boreal and temperate zones [13]. "*

R1-3. 190: occurrence OR abundance? which one when? Maybe add "(See Methods, section X)"

Author Response:

Use of occurrence or abundance is explained in the first paragraph of the methods, so we have added (see Methods) as suggested.

R1-4. 197: (also 819-821) This needs a little more explanation. I find myself struggling to understand the meaning of the blue and black error bars, since some of the blue ones are negative and some black ones are positive. I don't get it.

Author Response:

We have completely revised the figures and captions which makes it easier to understand what each component means. This included presenting mean negative effects, overall effects, and mean positive effects in separate plots, with the only vertical line at the zero-effect mark on the x axis.

R1-5. 201: "biodiversity" not defined, is this species richness? important for the interpretation of the effects

Author Response:

Referee 2 also identified this misuse of 'biodiversity'. In this paragraph we have changed the first use of biodiversity to the actual response metrics: abundance or occurrence. To avoid repetitive use of that phrase, we replaced the second occurrence with 'plant and animal responses'. We retained the third use of 'biodiversity' where it was used in a more general context and not in specific reference to our response variables. We could find no other instances where we used 'biodiversity' when referring to response variables. "Biodiversity" is used in the more general sense in the concluding sections.

R1-6. 215-216: This is pretty vague

Author Response:

We have reworded this sentence to make it clearer what we intended; more of a segue to the next section and explanation of why we do not discuss fire severity main effects in detail.

" We next consider how severity interacted with additional covariates, revealing that fire severity can have stronger effects in particular circumstances."

R1-7. 222-224: Were these frequently burned areas places that are adapted to frequent surface fire, and then burned at high severity (i.e. stand-replacing)? Might there be a reason to believe that places that hadn't burned were adapted to infrequent, stand-replacing fires? Low severity had positive or neutral effects on places with 3+ fires since '79, suggesting that areas with a high-frequency fire regime were less impacted when severity was low. High severity had neutral effects but overall negative on places with 0 fires since '79

Author Response:

This is a good point to consider: the idea that fire frequency might simply reflect ecosystem type, with rainforests representing the sites with very few past burns and grasslands representing the majority of sites with three or more burns.

We evaluated this risk by considering the proportion of sites burnt by 0 to 3+ fires since 1979 for each ecosystem group (Table below). We found that while there are some biases in the expected direction for most ecosystems, each ecosystem is nevertheless well represented across the range of number of fires.

Ecosystem Group	0	1	2	3+
Alpine-subalpine	0.12	0.41	0.4	0.07
Dry sclerophyll	0.38	0.32	0.15	0.15
Grass-heathland	0.22	0.37	0.21	0.19
Rainforest	0.51	0.28	0.13	0.08
Wet sclerophyll	0.4	0.3	0.18	0.12
Wetland	0.6	0.26	0.07	0.07

Consequently, we now acknowledge this issue in the second last paragraph of the "Effect sizes" section in the methods and have added the above table as Extended Data Table 4.

" The number of past fires was weakly related to the ecosystem group, with all ecosystems represented within each category of number of past fires (Extended Data Table 4). "

R1-8. 355 -65: probably can't draw too many conclusions about plants, since an assessment of the recovery of trees, shrubs and perennials might necessarily require 5 or more years after disturbance

Author Response:

We agree, and now remind readers in this sentence that our results relate to the specific period after fire for which our data were collected.

" Plants, insects, and reptiles had moderate negative and positive effects within the study period, regardless of fire severity "

We have also added a new paragraph (third last paragraph of the main text) that explains why our data from the first two years after the fires provide important insight into fire responses. (see quote in response to R2-10 below).

R1-9. Fig. 4: It seems odd to me that there are no grasslands? It seems like it might be good to include another landcover-based effect size estimation, with open-canopy ecosystems like grasslands and savannas versus closed-canopy systems.

Author Response:

See our response to part of major issue C/D above where we explain there are no savannas in the study area.

R1-10. 438: this study really is looking at resistance to fire, and resilience to fire is probably not able to be assessed from this data.

Author Response:

Our responses examine changes in abundance or occurrence after a disturbance, so we agree that these are measures of resistance rather than resilience. Resilience includes recovery after initial decline, and we have suggested that movement from unburnt areas or from surrounding protected areas can aid recovery after fire, providing resilience. We have carefully considered each use of the term resilience and replaced them with resistance where results refer to changes in abundance or occurrence. We have retained resilience where it is used in the context of movement that facilitates recovery.

R1-11. 466: I might add joint species distribution models

Author Response:

We have added a reference to a relevant example, but have also reduced the detail to streamline the sentence. The new sentence, in the second paragraph of the Future Research section, reads:

" When combined with spatial data including drought, other disturbances [14], fire history, and post-fire vegetation recovery, new modelling approaches[15,16] may allow for mechanistic prediction of plant and animal responses to fire. "

R1-12. 738: Typo: "unburntburnt"

Author Response:

In the version that we submitted this appeared as "unburntburnt". In the revision this sentence has been removed.

R1-13. 793-4: "test = "t"; dfs="contain"; method = "REML" 69. We also set control=list(optimizer="optim", optmethod="Nelder-Mead")". write out what all this means instead of copy and pasting R code. e.g. we used a T test and restricted maximum likelihood, with the Nelder-Mean method of model optimization. Make a github repository that fully reproduces the analysis and refer to that for folks that want to see the code.

Author Response:

We have reworded this sentence: *" We used restricted maximum likelihood[17], with the Nelder-Mean method of model optimization which enabled most models to converge. We used a T test with denominator degrees of freedom adjusted for random effects using the 'contain' option[17,18]."*

We have added our code and datasets to a Figshare repository; please see links in the data-availability statement. The full site level dataset is available for review (contact amy.irving@springernature.com, a.armstrong@nature.com).

R1-14. 808-813: it looks like you're doing some model selection. In that case, I think the proper way to do it is to use maximum likelihood to compare models, then use REML for the chosen model. This probably wouldn't change much, but it seems like it might be a good idea to make sure you do it like that.

Author Response:

We were not able to use model selection because each analysis uses a different dataset. We have now improved clarity in two sections of the methods to make this clearer

First, at the end of the effect sizes section, we now say:

" We calculated effects sizes for species/taxa within projects using an appropriate subset of sites. Effect sizes were calculated for unburnt vs burnt sites, and two fire-severity contrasts where there were enough sites burnt at that severity within the project: unburnt vs high/extreme severity, and unburnt vs low/medium severity (Extended Data Table 1). We did this using all sites within the project, and then using each subset of sites based on covariate values (Extended Data Fig. 6). For example, with sites experiencing one of three levels of pre-fire drought within a project, we calculated three effect sizes for species X, using unburnt sites compared with: (1) all burnt sites, (2) just the low severity sites, and (3) just the high severity sites (Extended Data Fig. 1). This was repeated for each level of drought."

Second, within the meta-analysis section we now say:

" Where there was no evidence for an interaction between a covariate and fire severity (P > 0.1) we applied a univariate model with the covariate as the only moderator. In these cases, we used the effects calculated using unburnt vs burnt sites rather than unburnt vs sites burnt at either high or low severity. "

R1-15. 834: Still do a github repository that has all the code from the analysis along with links publicly available data used, so that someone making reasonable request for the data can repeat the analysis.

Author Response:

We have created a Figshare repository. Please see links in the Data availability statement.

R1-16. general methods comment: might be good to italicize r package and function names.

Author Response:

We have italicized package and function names in the revision.

Referee #2 (Remarks to the Author):

This ms presents the results from a meta-analysis of impact studies related to the massive 2019-2020 fires in Australia. While it is based on a unique dataset and reveals potentially important findings, I found the presentation of the work to be problematic and less compelling than it could have been (IMO). I hope the following points are useful in crafting an improved revision.

Author Response:

Thanks for your careful evaluation of our manuscript. A substantial theme of R2 is about where effects take place. In our response to R2-6 below, we explain that our results are not spatially located, but are based on effects within projects. The projects have sites spanning hundreds to over 1000 km. Nevertheless, to provide some new insights into how effect sizes might be might vary across south-eastern and southern Australia, we have included a new covariate, Region (see R2-6 for full details).

R2-1) Abstract (and throughout): There is a general emphasis on the negative effects of fire, in the interpretation of findings, despite often reporting strong positive effects of fire. This is a bit confusing and at times feels like “cherry picking” results to support specific management and policy positions.

Author Response:

We have carefully considered each set of results and have added additional text to interpret positive effects where that had been omitted. We found two areas where additions were needed.

Section Unburnt area within 2.5 km: *" The importance of immigration from surrounding unburnt areas is also implied by the large positive effects associated with high levels of unburnt area (53%, 24% larger than mid and low sites respectively, Fig. 2d)"*

Section Taxonomic group, when discussing mammal responses: *" On the flip side, mammals may have the largest increases through an influx of wide-ranging species, particularly herbivores that eat fresh regrowth [19], and reduced competition associated with large declines by mammals."*

In our revision of the second paragraph of the Frequency and interval section, we explain more clearly that negative effects are of greatest conservation concern because fire is becoming more extreme with the changing climate, making habitats with mild fire regimes rare. Conversely, taxa that increase in more extreme fire regimes will tend to increase over the coming years, so will be of less conservation concern. This clarifies why we give more emphasis to patterns of decline.

" Habitat that is created and maintained through frequent and high severity fires is vital for some species [20] and it is becoming more common due to climate change [21]. Species that decline after fire are therefore of greatest conservation concern, especially in regions with increased fire activity. "

R2-2) Line 201: Beginning here the authors essentially refer to “biodiversity” as the dependent variable, when it is actually not species richness (or some other common metric of biodiversity) that is being examined. Instead, it is really reflecting occurrence or abundance of whatever plant or animal species happened to be part of the studies in the meta-analysis. I’m not sure what better term would fit here, but we are not dealing with “biodiversity” as it is usually used in scientific research and conversation.

Author Response:

Referee 1 also identified this mis-use of 'biodiversity'. In this paragraph (paragraph before the Fire severity section) we have changed the first use of biodiversity to the actual response metrics: abundance or occurrence. To avoid repetitive use of that phrase, we replaced the second occurrence with 'plant and animal responses'. We deleted the sentence with the third use of 'biodiversity' to move more quickly into the results. We could find no other instances where we used 'biodiversity' when referring to response variables. "Biodiversity" is used in the more general sense in the concluding sections.

R2-3) Line 212: The first sentence here is quite circular – "More severe fire is expected to have more severe ecological effects" – as "severity" is an ecological concept. It glosses over what characteristics of a given fire lead to more ecological impacts. For example, sometimes frequent fires are ecologically severe, while in other systems it's high intensity fires that have severe ecological impacts. In still other systems, it may be larger fires that are most severe. What fire characteristics are being examined? If we're talking about heat release or canopy consumption, it would be helpful to be more specific, without going deeply into the methods here.

Author Response:

We have reworded this sentence to avoid the impression of circularity. We clarify what is meant by severity by adding a reference to Extended Data Table 1. In addition to the two references that we gave as examples of where high severity fire had stronger effects on animals than low severity fire, we have replaced one reference with Keith et al. (2022), who recently discussed high severity fire as a threat to ecosystems.

" Severe fire, by definition, burns more plant biomass than low severity fire (Extended Data Table 1) and is therefore expected to have larger effects across all taxonomic groups [22,23]"

R2-4) Line 219-220: The first sentence about short interval fires is generally not true for grasslands, which can reburn very often, nor for many of the world's mixed conifer forests that (prehistorically) burned often at low severity (i.e., consuming understory and benefiting dominant overstory trees). HOWEVER, the notion of "fire deficit" in western US forests is so widespread that many will find this sentence to be problematic. More nuance/context needed for international readership.

Author Response:

We have revised this sentence to clarify that we mean fires more frequent than the historic range of frequencies. There will always be a frequency that is too high. Carlsen et al (2017) is an example where annual burning in Californian grassland has detrimental effects on plant species compared with burning every three years, supporting the generality of our revised statement. We have not added this citation to the list because we are over the recommended limit of 50 references.

"Fire that is too frequent can make plants vulnerable to decline by limiting the time available to set seed or replenish energy stores [24,25]"

R2-5) Line 228-230: Again, although potentially important and true in this context, it is in opposition to the "fire deficit" bandwagon mentioned earlier. More nuance/context needed for wide audience.

Author Response:

In the first paragraph of the Frequency and interval section, we have clarified that our statement here is in relation to Australian forests, in addition to other changes that improves alignment of our interpretation with our results.

"These effects were evident even for low-severity fires, which is of particular concern for fire management. It implies that even where prescribed burning in Australian forest has reduced the severity of a wildfire, large declines are still likely if past fires were recent or frequent."

R2-6) Line 232-240: This paragraph reports positive effects – seeming to contradict the previous paragraph – without the context needed to understand why they occurred. Further, the reporting of both positive and negative effects, for largely the same variables, happens repeatedly throughout the rest of the ms. You offer some literature and/or hypotheses as to why different effects might be observed, but this whole presentation is not very satisfying without a deeper understanding of WHERE the effects are being observed. Figure 2 is also increasingly frustrating and unsatisfying as a key way to show us what is going on.

Author Response:

Our fourth paragraph after the summary provides a brief overview of the methods. In that paragraph we explain:

"Meta-analysis typically focuses on the average mean effect of an intervention, but this masks responses in opposing directions. We therefore also report mean effect sizes for positive and negative effects separately, given the expectation that some species will increase, and others decline after fire via different mechanisms."

This makes it clear what we are reporting positive and negative mean effects, in addition to the overall mean effect. We have added a new figure, Fig. 2, to help readers understand that we are reporting the positive and negative effects. To further help readers with the figures, we have revised the layout to make the three components more obvious, and added additional clarification to the caption on what positive, negative, and overall mean effect sizes are:

" Left panels, black: negative mean effect size (i.e. the average effect size for all effects < 0); central panels, grey: overall mean effect sizes (i.e. mean of all negative, zero and positive effects): right panels, blue: positive mean effect sizes (i.e. the average effect size for all effects > 0)"

This comment and others further down by Referee 2 have a focus on where effects are being observed. However, effect sizes cannot easily be mapped because they are made up of multiple sites based on covariates of interest spanning up to 1631km.

We explained in the fourth paragraph of the main text addressing methods that:

" Using meta-analytical methods, we evaluated the effect of fire severity and its interaction with fire frequency, inter-fire interval, fire extent, drought, and protected areas (Extended Data Table 1), fitting project, species and broad taxonomic group as random effects (Methods). We did this by selecting subsets of sites within projects that represented specific levels of each covariate, then calculated effect sizes by comparing abundance or occurrence in unburnt sites with sites burnt at low or high severity (Extended Data Fig. 1)."

This overview gets across the key point that effects are calculated using multiple sites within projects. We have added text to the second paragraph of the methods to explain the spatial scale of projects:

"Individual projects had multiple sites spanning an average of 192 km (SD 284 km) with the most extensive study spanning 1631 km."

This now makes it clearer that mean effects (from the sites within projects) do not represent a specific location; they represent a mean effect across sites spanning up to 1631 km.

Nevertheless, the desire to understand spatial differences in effect sizes can be represented to some extent using our methods. We have created a new covariate called 'region' which divides the study area into three regions based on latitude (see revised Extended Data Table 1). We provide these results in Extended Data Fig. 2, provide a short interpretation within the figure caption, and mention it briefly in the first paragraph of the methods section to avoid further complicating the manuscript.

Extended data Fig. 2 caption now includes: *" The most striking regional effect was for weaker positive effects in the south, particularly at high severity (d). Fire conditions differed across regions. Our data suggest a weak trend towards stronger pre-fire drought in the south (Pearson's correlation with latitude = -0.03, t = -2.5343, df = 6239, P = 0.01). Fires that drove the world's most extreme pyrocumulonimbus event were all in our mid and southern regions, corresponding with larger contiguous areas of high severity fire [26]. Remote sensing data indicate that the southern parts of the study region had slower rates of vegetation recovery than in the north [27]. These differences in pre-fire conditions, fire behaviour, and recovery rates suggest that the level of disturbance and other stressors can constrain the response of taxa which would otherwise benefit from bushfires."*

In the first paragraph of the methods: *" Effect sizes showed some regional variation (Extended Data Fig. 2d)."*

R2-7) Lines 243-245: The authors assert that “Conservation and management in fire-prone regions must therefore focus on maintaining areas with intermediate inter-fire intervals, where we observed the smallest effect sizes (Fig. 2c)…” This seems counter-intuitive. Why wouldn’t you instead aim for the fire intervals (and severities) that coincided with the biggest positive effects on biodiversity, in those places where that was relevant. In contrast, in environments where a certain type of fire had big negative effects, youpos would avoid using that kind of fire regime. Why aim for the smallest effect sizes? It seems like one of the important insights of this study is to identify which types of fire regimes are best for biodiversity in which parts of the study area… Again, the spatial aspect is missing.

Author Response:

As explained in response to R2-6 above, there is no specific location of positive and negative effects. There are both positive and negative mean effects, based on means from sites within projects that span hundreds of kilometres.

However, the question of why focus on smaller effects is a good one, and we did not explain this clearly. We have expanded this paragraph to make our logic explicit.

" Habitat that is created and maintained through frequent and high severity fires is vital for some species [20] and it is becoming more common due to climate change [21]. Species that

decline after fire are therefore of greatest conservation concern, especially in regions with increased fire activity. A key management objective will be to maintain areas with intermediate and long inter-fire intervals. We observed the smallest effect sizes at intermediate fire intervals (11-20 years, Fig. 3c), indicating that communities undergo the least disruption at these intervals. Long intervals are also needed to serve as refuges [28], create time-dependent habitat attributes such as tree hollows[29], and support source populations for species that might be lost from areas burnt too frequently [30]."

R2-8) The effects figures quite difficult to read and absorb, including the interpretation of P values. Are all of the "Reduced" and "Increased" significantly different from zero, while only a few of the "Overall" are significant? What does it mean when the "Severity P" is not significant in some panels?

Author Response:

We have revised the figures to make the information presented more intuitive. The changes make it clear that the P values refer to either differences among levels of the covariates, or interactions between covariates and fire severity.

We used an overall test for an effect of severity (Fig. 2a), but main effects of severity were also tested in each model as part of the analysis structure, even though main effects of severity were not of interest in the covariate analyses. In our revision, only the pertinent P values are provided on the figures; either the main effect P value or the P value for the interaction with severity. The P values for the main effects of severity in models testing for interactions between covariates and severity are still available in Supplementary Information Table 1.

The text does not seem to differentiate by significance when reporting results… why?

Author Response:

This question highlighted a weakness of our original presentation. We provide P values for the main effect and the interaction with fire severity on the original figures, but we always plotted the interaction if any one of the negative, positive, or overall mean effects had a significant interaction. Then in the text we reported mean effects for the interaction, even when a particular interaction was not significant, which was not appropriate.

In our revised manuscript we present main effects in the figures when the interaction was not significant and talk about the main effects in the text. For example, in Fig. 3b, negative effects and positive effects did not have a significant interaction and so main effects are displayed and mentioned in the text. The row labels make it clear what is displayed (an interaction or main effect), along with guiding text in the revised caption.

These new figures include P values for the result displayed (either main effect or interaction), so there is, in most cases, no need to repeat the P values in the main text. The full statistics including the test statistic, degrees of freedom and P value are available in Supplementary Information Table 1. In our revision, our text does not report trends that do not have statistical support.

To bring reader's attention to results that have weak statistical support, we have edited the text, while also trying to avoid making the main text cumbersome. In one instance we have provided additional information about P values in the caption for Fig. 4. *" Although the interaction of severity with drought for negative effects was weakly supported (a), the main effect was strongly supported (P = 0.01, Supplementary Information Table 1)."*

In another example, in the first paragraph of Protected areas section, we now point out: *"Our results show that unprotected areas had 23% larger negative and 26% larger positive effect sizes than protected areas, although statistical support was weak (Fig. 4b) and absolute changes were small (2% Extended Data Table 2)."*

Are the "Overall" effects ever really meaningful, and are they worth reporting? It's where the positive (or negative) effects occur and why that is interesting, no?

Author Response:

We agree on this point; the positive and negative mean effects allow more substantive understanding of the ecology of responses to fire. That is why we give little emphasis to the overall effects in the text. However, up to now, meta-analysis studies have only reported the overall mean effect. We feel it is important to report the overall mean effect so that people can see the different insights that each summary statistic provides. Also, it provides a handy indicator of how the balance between positive and negative effects played out, even though the ecological insights we can draw from that are limited.

R2-9) Lines 289-299: Again, it is confusing and causes a bit of "whiplash" as one finishes a relatively compelling paragraph on negative effects (on plants? animals?), but then is presented by largely contradictory results in this new paragraph.

Author Response:

We have explained the nature of negative and positive effects in response to R2-6 above. We hope that clarifies the logic of discussing negative responses and positive responses in sequential paragraphs.

We have edited the fourth paragraph which briefly outlines the methods, to clarify how taxonomic groups are handled, a point which is further detailed in the methods section.

" Using meta-analytical methods, we evaluated the effect of fire severity and its interaction with fire frequency, inter-fire interval, fire extent, drought, and protected areas (Extended Data Table 1), fitting project, species and broad taxonomic group as random effects (Methods). "

R2-10) Lines 294-299: While some of these hypotheses are interesting, we are looking at a very recent fire and only ~3 yr of recovery. It seems very unlikely that these results are reflective of longer term recovery patterns, making their relevance questionable for wider management and conservation questions.

Author Response:

This is an important point, and we agree that more explanation is warranted. We have divided our original second last paragraph, expanding on the time-frame issue in a new stand-alone paragraph. This new paragraph explains the importance of changes in the short term and how the often large and abrupt changes associated with a fire event have long-term implications. By implication, understanding the mechanisms driving these abrupt short-term changes is important for management, as detailed in the Future Management section.

" Although our data capture short term responses, they are by no means trivial. Time-since-fire response curves [31,32] show responses to fire are most acute in the months and years

immediately following fire. Our data capture these abrupt changes, and show how the magnitude of change can be moderated by covariates, particularly fire history. While post-fire succession continues for decades or centuries, for many species the abrupt short-term changes [33], or their absence [34], have a strong influence on subsequent successional pathways. Further, our dataset includes over 2200 taxa, likely including species with diverse long-term responses. Some species are expected to return to pre-fire numbers within a few years, with no long-lasting conservation implications. Others that declined may struggle to recover [35] or may never return [36]. Our results therefore offer a useful guide for conservation action. However, investment in widespread long-term monitoring is needed to identify species that may not have time to recover before the next fires. "

R2-11) Lines 302-305: This seems like a prime opportunity to show WHERE the positive vs negative effects were observed, no? Maybe I'm not understanding something about the analyses, but the sites are shown in Fig 1, so mapping out the results would help greatly with interpretation and could also reveal additional variables that could be included in the analysis.

Author Response:

See our response to R2-6 above where we explain that the mean effects are not spatially located, but we have added a new analysis to report broad geographic trends.

R2-12) Line 305: The "buffering effect" of protected areas highlighted here and later on is an unusual interpretation. Granted, effects are generally less in protected areas, but they are also not experiencing the 26% positive effect in some areas either. Is this a good thing? Where the positive effects are happening (and to which taxa) is important information, and knowing this would impact the entire ms.

Author Response:

As for our response to R2-6, the mean effects are not spatially located, although our supplementary analysis by region suggests some north-south bias in positive effect sizes for sites burnt at high severity (Extended data Figure 1). We explain in our revised Frequency and Interval section, third paragraph, that the negative effects are of most conservation concern (see quote in response to R2-7). That explanation comes earlier in the discussion than the buffering effect of protected areas, providing appropriate context to understand why the buffering effect is valuable from a conservation perspective.

We also now reiterate that point about decliners being of conservation concern in our discussion of protected areas. *" The apparent buffering effect of protected areas could deliver overall conservation gains by minimising declines. Although this will also lead to smaller increases after fire, taxa that increase are likely at least conservation risk because burnt area is becoming more common [21]."*

R2-13) Lines 248-253: How do these results compare with other longer-term studies across animal taxa? For example:
https://onlinelibrary.wiley.com/doi/full/10.1111/ele.14177

Author Response:

The suggested study would take a new paragraph to appropriately contextualise. Moritz et al. [37] used the proportion of net primary productivity consumed by fire at a course resolution (0.25^0) as a fire index, considering the entire globe. They therefore include savannas and grasslands that have very high values for their fire index, but likely have low intensity and so mammals may be less impacted than similarly high values in a forest. Their biodiversity response was species richness based on distribution data and so their responses are related to biogeographic and evolutionary history; they indicate in their discussion that they cannot distinguish evolutionary from ecological explanations.

Rather than expanding on longer term responses of mammals in this section, we have instead addressed the importance of our short-term results in a new, third-last paragraph of the discussion, as described in response to R2-10 above.

We have also provided additional international context in the paragraph about mammals by referring to a recent meta-analysis that reports mammals are more sensitive than other taxa. The start of the paragraph now reads:

" Mammals had larger positive and negative effect sizes than all other taxa, especially at sites burnt at high severity where negative effects were 94-155% larger and positive effects 66-96% larger than other taxa (Fig. 5a). This is consistent with a global meta-analysis where mammal abundance and richness had negative effects while other taxa had neutral effects[38]."

R2-14) Lines 400-411: You emphasize the negative effects here, but at nearly every turn you present substantial contrasting negative AND positive effects.

Author Response:

Please see our response to this problem in R2-6 above.

Is there a reason to discard most of the positive effects and focus strongly just on the negative? Could you make the case that the positive effects are fleeting or artifacts and thus less useful? Can't you use the location of the sites to better inform a mapped interpretation? (Sorry, broken record...)

Author Response:

We explain why we give an emphasis to negative effects in our response to R2-7 above. But we have also carefully considered where such a bias was undue, and have added a more balanced consideration of negative and positive results. These include in the protected areas section (see quote in R2-12), and in the paragraph about mammal responses in the taxonomic group section of the discussion (see quote in R2-1).

R2-15) Lines 432-435: Where are these refuges within the study area, and how did they fare?

Author Response:

We explain in response to R2-6 above that the results are not spatial. Although it would be possible to use the spatial drought layer to find locations with low or high pre-fire drought, a thorough evaluation of those data is beyond the scope of this study. To make a useful contribution, temporal stability of refuges would need to be evaluated; it would be a whole separate project.

R2-16) Table 1: This table is useful for the reader, but it feels a little late in the game to make sense of all of the preceding (apparently) contradictory results. It could help to give a shorter overview of such mechanisms earlier in the paper.

Author Response:

We now also refer to Table 1 in the earlier Frequency and Interval section, bringing reader's attention to this handy summary of mechanisms.

Why isn't a table like this based on the large literature that exists on how fire interacts (positively and negatively) with plants and animals?

Bond & Keeley, 2005; Carbone et al., 2019; He et al., 2019; Keeley et al., 2011; McLauchlan et al., 2020; Moritz et al, 2023; Nimmo et al., 2019; Pausas & Parr, 2018; Santos et al, 2023

Author Response:

The table is a summary of our results, based on 62 projects where we had site-level data from the 2019-20 megafires. We did not systematically review other studies so other studies are not part of our results.

R2-17) Line 457 onward: The proposed mechanisms could be supported by, or even generated by, mapping out where you found evidence of such pathways, based on the sites in the study and the effects observed. This would be a much more compelling way to justify the mechanisms presented in this table. I would suggest adding such an analysis, even if relatively superficial.

Author Response:

As discussed in our response to R2-6 above, the results are not spatially locatable.

Referee #3 (Remarks to the Author):

My review is focused on the dataset and the meta-analysis.

Dataset

R3-1- The authors performed a meta-analysis in order to estimate impact of mega-fires on biodiversity. However, their meta-analysis does not seem to be based on a systematic review of the available studies but to a set of "projets" whose origin is a bit unclear. Although the list of projets is provided, the authors did not explain how these projects were retrieved. The authors should have implemented a systematic search strategy and described it in details in order to ensure transparency and reproducibility. They should explain how their search strategy was designed to ensure that all relevant projects were retrieved.

Author Response:

You are correct that we have not systematically reviewed the literature for this project. A systematic review would not have met our goals because most datasets from the 2019-20 fires are not published. Therefore, although we use meta-analytical models, this is not a traditional meta-analysis using data extracted from the published literature. Instead, this is a large collaboration based on the 2019-20 fires whereby authors have contributed their own site-level datasets to the project. This is a strength of the study beyond the standard systematic review and meta-analysis. All studies relate to the 2019-20 fire events, and so our data have less variability than is inherent in many ecological meta-analyses. Another key strength is that by having site-level data, we can sample sites within projects to ask specific questions about covariates that the individual study is not capable of addressing, because individual studies do not span a large enough range of covariate values. Therefore, we have been able to explore a range of important covariates that would not have appeared if using a systematic review of published literature.

We have revised the main text, third paragraph after the summary, to help readers understand this at the opening of the manuscript: *"Here, we collate site-level data from 62 datasets with over 810,000 abundance, presence, or absence records, including over 2200 taxa…"*.

We have also revised the first paragraph of the methods to provide more detail about how the projects were gathered.

" We refer to megafires as those >100,000ha, with 12 megafires recognised throughout the 2019–20 fire area[39], although proposed terminology would classify these as giga- and tera-fires [40]. We gathered site-level data for 62 projects (Supplementary information Table 2) by inviting contributors to join our collaboration. We emailed leaders of projects studying post-fire biodiversity impacts that were funded by the Australian Government's Wildlife and Habitat Bushfire Recovery Program. We promoted our proposal on social media in summer 2021-2, ran a workshop at the Ecological Society of Australia conference in November 2021 and held a two-day on-line symposium in February 2022. Projects were collated throughout 2022 after potential contributors contacted us. "

R3-2- In addition, I was wondering why the authors focused in Australia only, while mega-fires occured in other areas as well (e.g. California). Because of this choice, it is difficult to understand how their results could be extrapolated.

Author Response:

The Australian 2019-20 fires were, at the time, the most extensive fires globally, and remain significant in their own right. For example, the 2020-21 Californian fires burned 19000 km^2 [41], whereas the Australian fires burned over five times that area. In fact, the area burned as high severity, assuming crown scorch or consumption, during the Australian fires was almost as large as the entire area burned in the Californian 2020-21 fires (16,000 km^2). We did not undertake a systematic review of already published datasets, which might easily take a global perspective. Instead, we carefully assembled a network of collaborators, and gathered a very large collection of unpublished datasets that were then consolidated into a common format and structure, and we obtained new covariates across the range of sites and projects. To our knowledge, nobody has undertaken such an effort to date.

The scope of our dataset spans all vertebrate groups, insects, molluscs, and plants, and covers an extensive range of forested climates, from mediterranean in SW Western Australia, to cool temperate, warm temperate and subtropical climates along the Australian east coast. Given this diverse range of taxa, ecosystems, and regions, it seems reasonable to assume general patterns and valuable lessons have emerged for other fire-prone forested landscapes around the world. Importantly, our study raises several challenging hypotheses that countries experiencing megafires will need to test to understand the nuances of differences among countries.

In response to Referee 4-8 and R1-1, we have expanded our discussion of management implications to better contextualise the work using management-related research in other parts of the world.

R3-3- The authors listed a series of selection criteria (lines 710-719) but did not justify them very clearly. Moreover, they did not explain how the data were extracted.

Author Response:

We have revised the first paragraph of the methods to better explain why we used each selection criterion.

" Projects had to include site level data from sites burnt in 2019–20 within the megafire region [1]. To ensure a minimum level of treatment replication, enabling means and standard deviations to be estimated, datasets needed at least three burnt and unburnt sites (e.g., survey points, plots, or transects), or three sites with beforeafter data and at least three non-zero records across all sites (at the species/taxon level). Projects also needed to report either an estimate or index of abundance (44 projects), occurrence (17 projects), or probability of occurrence after accounting for detection (1 project). Although we canvassed researchers for a full range of response variables, the datasets provided included almost exclusively presence-absence or abundance data. Abundance and occurrence effect sizes were similar, justifying using both together (Extended Data Fig. 2c). "

R3-4- It would be useful to produce a diagram similar to the PRISMA diagram (a standard diagram in systematic review and meta-analysis) in order to present the results of their review and selection procedure.

Author Response:

As clarified above (R3-1) this study is not a systematic review, and as such a PRISMA diagram is not appropriate here. Nevertheless, we have now added mention of a small number of projects that were offered but did not meet our criteria to provide transparency in how projects were obtained.

In the first paragraph of the methods section:

" *Only one project met all other criteria but reported responses that were not related to abundance or occurrence (effects of grazing on reproduction in one orchid species), and thus was omitted from analysis. Five other projects were excluded because they did not have at least three unburnt sites.*"

R3-5- More generally, I advise the authors to consult standard guidelines on systematic review.

Author Response:

The changes made in response to the previous four comments will ensure readers of the revised manuscript understand that our study is not a systematic review.

R3-6- Effect size and meta-analysis

- The authors used different types of biodiversity measures and they mixed occurence and abundance of taxa together. It would be interesting to check whether the results are different between these two types of metrics.

Author Response:

We have reanalysed the data to include evaluation of how occurrence and abundance effect sizes differ, and whether it interacts with fire severity. The results are included in Extended Data Fig. 2, and show only minor differences between these response types. We have explained in the caption to that revised figure: *" Abundance (Abund) and occurrence (Occur) responses showed very similar effect sizes, with the only exception of higher positive effect sizes for occurrence data at high severity. The small number of effects in this group precludes further separate consideration of occurrence responses (c). "*

We draw reader's attention to this in the first paragraph of the methods: *"Abundance and occurrence effect sizes were similar, justifying using both together (Extended data Fig. 2)."*

R3-7- 723: do you mean « one to five » ?

Author Response:

We have edited this to read "one to five".

R3-8- I don't fully understand why the authors used standardized mean difference (SMD) and not (log) ratio as an effect size. The problem with SMD is that the differences between groups may be due to either difference between mean values or difference between standard deviations. This makes the interpretation of the results a bit difficult.

Author Response:

We have given this careful thought and now provide a new approach, which we explain in the effect sizes section of the methods (see large quote below). Our use of mean change effect sizes generally supports the standardised effect sizes and provides additional insight into the magnitude of effects, which we now incorporate into our improved interpretation using a new summary in Extended Data Table 2. Mean difference effect sizes are calculated as the difference between two values, and so values can be obtained even when the burnt or unburnt value is zero, unlike the log response ratio. This makes using mean differences more appropriate than log response ratios. We justify our use of mean difference effect sizes in the revised methods by noting that 21.6% of effect sizes included zero values for either the burnt or unburnt response.

In the methods:

" We used standardised mean change and standardised mean difference because these effect sizes are more generalisable than unstandardised effect sizes; they are more likely to reflect the true values of a larger pool of samples [42]. SMCC and SMD return NA values when the standard deviation is zero. A very small percentage of effect sizes (0.9%) had no variation, such as when a taxon was present on all sites before and after the fire. Omitting such values has minimal impact on our results because there are so few of them and because we focus on the size of negative and positive effects separately, rather than pooling all effect sizes including zeros (see Meta-analysis below). However, standardised effects can be difficult to interpret, and Takeshima et al. [42] recommend providing supplementary analyses using mean differences. We therefore also calculate effect sizes

using the MC option of escalc to calculate the raw mean change in before-after designs, and the MD option to calculate the raw mean difference for control-impact and BACI designs [18,43]. These gave very similar results to the standardised effect sizes (Extended Data Fig. 3, Supplementary Information Table 3), but provide additional insight into the scale of changes observed. Mean differences are the most appropriate effect size to use with our dataset where 21.6% of effect sizes included zero values for either the burnt or unburnt response.".

" Before using the MC and MD functions, data for species within projects were standardized by dividing all values by the maximum value for that species within that project, so that the data for calculating each effect size ranged from zero to one. While this makes data with different ranges of values comparable, it also risks biasing down effect sizes when there are large outliers. We therefore also calculate MC and MD after removing outliers. We assumed our count data (x) approximate a Poisson distribution, and by transforming our data as $Y = 2\sqrt{x}$, outliers can be identified as those greater than (median(Y)+3), based on the approximation that Y is normally distributed with a standard deviation of one [44]. This approach is conservative, such that if the Poisson lambda value is less than three, it will identify fewer outliers. When applied to binomial data, it does not identify any outliers. This produced results similar to when outliers were included, but effect sizes were slightly larger (Extended Data Fig. 4, Supplementary Information Table 4).

" Some variances for mean change and mean difference effects were zero. These prevented meta-analysis models from converging. To enable models to converge, and to use a dataset similar to the SMCC and SMD, we removed effects with zero variance before analysis (mean removed = 2.4% SD = 2.3%; Supplementary information Table 1)."

We also add a short sentence in the methods paragraph, fourth paragraph after the summary, to alert readers to the two types of effect size that we will be reporting.

" We used standardised effect sizes to maximise generality but supplemented this with mean difference effect sizes for insights into absolute differences in the response [42]. "

R3-9- In before-after studies, the authors used the standard deviation « before » and not a pooled estimate of SD. Why?

Author Response:

Thanks for raising this. Morris and Deshon[45] explain that in the case of the single-group pre-test/post-test design (ie: a before-after design with repeated measures), differences should be standardized using the variability of change scores. Therefore, in repeating the analysis, we have used SMCC for the before-after designs, instead of our original SMCR. SMCC standardises using the variability of change scores.

R3-10- How the differences in sample sizes were taken into account? Maybe all samples had the same sizes, but this is not fully clear.

Author Response:

We assume the referee is referring to the number of sites used to calculate each effect size. Those sample sizes are accommodated within the analysis. When calculating effect sizes from original data, sample size is included in the function. When calculating mean effects within the meta-analysis, the analysis accommodates the estimated variance for each effect, which is a function of sample size.

R3-11- The standard deviations were computed using different methods, depending on the protocols of the projects. It is unclear why another method was used for control-impact studies. Why not using the same method for all studies?

Author Response:

The details of the escalc function from the metafor package [18] (https://wviechtb.github.io/metafor/reference/escalc.html?q=smcc#details), also Morris and Deshun [45] indicate that particular effect sizes are appropriate for different study designs. Due to the non-independence of before-after studies, they must be calculated using a method that accommodates that non-independence, such as using the MC or SMCC options. In contrast, the burnt-unburnt designs have independent sites in each treatment and must be calculated using MD or SMD.

We have revised the methods text (first paragraph of effect sizes section) to make it clearer why the SMCC effect size is needed for before-after studies.

" Effect sizes for these datasets were calculated as the standardized mean change (mean-after minus mean-before) using change score standardization (SMCC option in escalc function of R package metafor) [18,43,45] *which accommodates the expected non-independence of repeated measures of the same sites*[45]*."*

R3-12- The authors excluded the zero data (no difference between burnt and unburnt). This can lead to a biased estimation (overestimation of the effect). The authors mentioned convergence problems, but this is not a good reason for removing data. In case of numerical problems, the authors could either add a small constant to the data or used a different statistical model able to deal with zero data (or use both strategies and compare the results). At least, the authors should estimate the probabilities of zero with a separate statistical model (that may even include covariates, e.g., glmm). Then, the authors could either present the results of the estimated probabilities separately, or combine these estimated probabilities of zero with the mean effect size estimated from the non zero SMD data in order to derive a global average.

Author Response:

We did not exclude any zero effect sizes where the SD could be estimated. Effect sizes with zero variance were those where we applied an ad-hoc solution in an attempt to keep zeros in the analysis. However, the number of zero-variance values in the dataset is small relative to the number of effects in each analysis (2.4%, see R3-8), so their exclusion is unlikely to affect the outcome. More importantly, most of our interpretation, and the key results of interest, are for negative effects and positive effects. These exclude all zero values. Given the low importance of zero-variance results, we have decided that it is not justifiable to use ad-hoc methods to try to include them, when that ultimately requires further ad-hoc solutions to try to accommodate them in the analysis, all of which distracts from the main findings. Instead, we use the standard practice of calculating SMCC and SMD with the escalc function, which creates NA values where the variance cannot be calculated, and these are later excluded from analysis. Our response to R3-8 includes the revised methods text.

R3-13- Concerning the control-impact studies, it is not explained how the burnt and unburnt sites were paired together.

Author Response:

Unburnt and burnt sites are not paired, they are independent. We now clarify this in the revised text: *" Control-impact was the second design. These surveyed independent unburnt and burnt sites after the fires (n = 46)."*

R3-14- I didn't understand how the SD values were computed for the third type of design (756-748).

Author Response:

We have reworded this section using shorter sentences which helps clarify what we did.

" The third design included before¬after-control-impact (BACI) projects (n = 3). These were converted to control (unburnt) and impact (burnt) designs by subtracting the before-fire response from the after-fire response for each site[46]. Subsequently, the SMD was calculated using the same method as the second design."

R3-15- 762-763: as one single control was paired with two different treatments, the two resulting effect sizes were not independent. However, the authors took this non-independence into account. This is a good point.

Author Response:

We agree. Thank you.

R3-16- 826-832: Here, I understand that the authors used another selection criteria (risk of confounding with drought). This should be explained before, where the authors described the selection procedure. Moreover, it is unclear what the authors mean when they write « we found some evidence ». The authors should use a transparent rule for study selection.

Author Response:

We have shifted this information further up before the meta-analysis is explained so that the main dataset going into the meta-analysis is more easily understood. We have edited the text of this paragraph to explain what evidence we found. This includes referring to new Extended Data figures and Supplementary Information with results of repeating the analysis with before-after designs retained.

" In another initial step, we fitted a two-level factor delimiting before¬after survey designs from those that sampled burnt and unburnt sites after the fires. Before¬after designs were confounded by the breaking of a multi-year drought shortly after the bushfires. This meant that measured effects of the bushfires could be smaller due to compensating effects of higher rainfall. We found some evidence for this expectation, with mean effects shifted to the right for before-after compared with burnt-unburnt designs (Extended Data Fig. 2a). We therefore omitted before¬after designs from all other analyses. This removed 14 before¬after projects representing 6.4% of effect sizes and made no substantive difference to the outcomes of the analysis (Extended Data Fig. 5, Supplementary Information Table 7). Rejecting before-after designs in meta-analysis has been recommended due to the risk of confounding [47]."

R3-17- It would be useful to check the influence of the data with risk of confounding on the results and check the robustness of the conclusions. I would thus redo the analysis with this group data and compare the results with those obtained without. Alternatively, you can add a covariate related to the risk of drought in your statistical model in order to deal with this issue.

Author Response:

As described in response to the previous comment, we repeated the analysis and provide graphs in Extended Data Fig.s 3, 4 and 5, and detailed analysis output as Supplementary Information tables 3, 4 and 7.

Referee #4 (Remarks to the Author):

R4-1. Overall, this paper represents an impressively large-scale and somewhat standardized collection of ecological data examining the effects of the 2019-2020 Australian fires on biodiversity. The information contained in this paper has the potential to be uniquely influential scientific knowledge that will be a critically important contribution to the literature. The methods used are relatively standard, non-controversial meta-analytic approaches, and seem to be implemented correctly.

Author Response:

Thank you.

R4-2. A consequence of the ambition of the study is that a staggering breadth of empirical results are included in the paper, which is both useful but also overwhelming to the reader. The paper is challenging to read and follow simply because of the sheer number of comparisons and tests that are examined. As such, the paper appears as scattered and unfocused. Of course, I recognize that distilling the paper any further than it has already been distilled will be a challenge.

Author Response:

In response to this and similar comments from other referees about the complexity of the figures, we have re-designed the figures to simplify their presentation. This has removed some of the P values and separated the negative, overall and positive effects, which makes it easier to digest the results.

We agree that some further reduction and simplification of the story would be beneficial. We have therefore designed a new conceptual figure (Figure 2 in the revision) and placed it early in the manuscript so that readers can more quickly understand the main points about how higher levels of disturbance lead to larger positive and negative effect sizes.

R4-3. The authors idea to split up negative and positive effects was a good once, so as to avoid incorrect inferences of “no effect” when the individual effects are highly variable and include many positive and negative effects. However, a consequence of this decision is that the communication of effects becomes highly convoluted. For example, beginning at line 318: “Mean negative and positive effects were larger at high compared with low seRverity for Amphibians…” There are so many qualifying words in these types of statements that it is really challenging to follow for a reader.

Author Response:

We have re-worded this example to limit convolution: *"We found larger positive and negative effects at high severity sites for amphibians…."*

We have carefully read the rest of the manuscript and reworded other sentences with this kind of complexity to make the presentation easier to read. This includes swapping out 'negative mean effect size' and 'positive mean effect size' for 'decline' and 'increase' respectively. In some cases, we retain reference to 'effect size', but have removed 'mean' to simplify the language.

R4-4. Perhaps another part of the difficulty in tracking and interpreting the results is they are visualized in three very busy and similar-looking figures, each of which contain many rows of numbers, labels, colors, abbreviations, nested components, p-values, etc etc. I do not envy the challenge the authors face in visualizing the results, which again, are staggeringly extensive. The effect of the presentation of many dozens of tests and similar figures is that the paper is highly repetitive, and lacking in terms of synthetic take-home messages.

I would encourage the editors and authors to work together to try and find a path forward that allows all of this important information to be accessible to readers while providing a more synthetic, focused narrative in the main text that more effectively communicates key results to readers.

Author Response:

Also in response to R2-8, we have substantially revised the figures including removing much of the text and dividing the components so they are visually more distinct. As mentioned in response to R4-2 above, we have also created a new conceptual model that synthesizes a key part of the story and will assist readers to understand the positive and negative effects in our main results figures.

Specific comments:
R4-5. 145: “Fire management should aim to defend…” This statement comes across as an advocacy position. Could be rephrased to be more stance-neutral: “Fire management would most effectively conserve biodiversity by safeguarding fire-sensitive ecosystems…” (for example)

Author Response:

We have re-worded these last three sentences of the summary to better reflect the strength of effects of the key covariates. The text is still relatively frank about the most important management approach to mitigate impacts, which is important for highlighting the real-world implications of our research.

" Building wildfire resilience into ecosystems depends on reducing fire recurrence, which may require rapid wildfire suppression, especially in areas recently or frequently burnt. Defending fire-sensitive wet ecosystems, expanding protected areas, and considering localised drought levels can also contribute. These are urgent countermeasures to mitigate the impacts of more frequent megafires caused by anthropogenic climate change."

R4-6. 157: “…on record, generated a record…” consider rephrasing

Author Response:

We reworded to: *"The 2019–2020 fires in Australia (hereafter, the ‘2019–20 fires’) burnt 10.3 M ha* [39]*, including the largest documented area burnt at high severity* [48]*. They generated a record number of pyro-convective storms and disrupted the planet’s stratosphere* [26]*."*

R4-7. 249: How was the 2.5 km buffer size selected?

Author Response:

This is detailed in Extended Data Table 1, where we explain for unburnt area within 2.5 km: *"Fire extent mapping was used to delineate the proportion of unburnt area compared to total area (burnt and unburnt) within a 2.5 km radius from a given survey site. Buffer radii from 250 m 2.5 km were highly correlated (Pearson correlation coefficient > 0.86) so we use 2.5 km for consistency with the protected area buffer size."*

For protected areas within 2.5 km we explain: *"ACLUMv8 Landuse for % natural environments (undisturbed) within a 2.5 km radius around survey site. Buffers of increasing radii difference exhibited decreasing correlation, so we use our largest radius, 2.5 km, to minimise correlation with P.A or not (Pearson correlation 0.6). 2.5 km was an arbitrary choice, likely with little consequence given high correlations with other radii (e.g. 1 km, r = 0.94)."*

R4-8. 399+: In these section, I would love to see a few more linkages to how implications, fire management recommendations, etc., might be similar or differ for other systems. The effects presented here, while extensive, are limited to one region of Australia. Some or many of the recommendations will not transfer to other systems. It's important that this is discussed at least briefly.

Author Response:

Also in response to R1-1, we have added a new second paragraph to the future management section that explicitly addresses where fire suppression may be a helpful management approach.

" Increasing use of rapid wildfire suppression is likely to apply to other regions of the world that experience vast fires such as Canada in 2023 [2], and in ecosystems that are threatened by fire, such as rainforest[3]. Some forest types, like Australia's ash forest[4] or North America's coastal temperate forest[5] could benefit from rapid wildfire suppression continued over many decades or centuries, consistent with historic fire regimes. But this will not be the case everywhere. Some dry forests of western United States can be restored

by adding frequent, low severity fire to maintain an early successional understorey [5]. Further, many USA forests can benefit from allowing mixed severity wildfires to burn [6]. With increasingly severe fire weather, a range of counter measures are needed, tailored to specific ecosystems. Indigenous fire practices, placing the right fire regimes into the right country [7], is therefore increasingly important for improving fire management for biodiversity conservation [8]. Complementing these ecosystem-specific approaches, we expect rapid wildfire suppression will need to be used more often to limit the extent of land burnt at short intervals."

References

1 Ward, M. *et al.* Impact of 2019-2020 mega-fires on Australian fauna habitat. *Nature Ecology & Evolution* **4**, 1321-+, doi:10.1038/s41559-020-1251-1 (2020).

2 CIFFC. *Canadian interagency forest fire centre, Fire statistics.*, <https://ciffc.net/statistics> (2023).

3 Feng, X. *et al.* How deregulation, drought and increasing fire impact Amazonian biodiversity. *Nature* **597**, 516-+, doi:10.1038/s41586-021-03876-7 (2021).

4 Lindenmayer, D., Taylor, C., Bowd, E. & Zylstra, P. What did it used to look like? A case study from tall, wet mainland Mountain Ash forests prior to British invasion. *Austral Ecol.* **49**, e13520, doi:https://doi.org/10.1111/aec.13520 (2024).

5 Noss, R. F., Franklin, J. F., Baker, W. L., Schoennagel, T. & Moyle, P. B. Managing fire-prone forests in the western United States. *Front. Ecol. Environ.* **4**, 481-487 (2006).

6 DellaSala, D. A., Baker, B. C., Hanson, C. T., Ruediger, L. & Baker, W. Have western USA fire suppression and megafire active management approaches become a contemporary Sisyphus? *Biol. Conserv.* **268**, doi:10.1016/j.biocon.2022.109499 (2022).

7 Prober, S. M., Yuen, E., O'Connor, M. H. & Schultz, L. Ngadju kala: Australian Aboriginal fire knowledge in the Great Western Woodlands. *Austral Ecol.* **41**, 716-732, doi:https://doi.org/10.1111/aec.12377 (2016).

8 Hoffman, K. M. *et al.* Conservation of Earth's biodiversity is embedded in Indigenous fire stewardship. *Proc. Natl. Acad. Sci. USA.* **118**, 6, doi:10.1073/pnas.2105073118 (2021).

9 Ellis, T. M., Bowman, D. M., Jain, P., Flannigan, M. D. & Williamson, G. J. Global increase in wildfire risk due to climate-driven declines in fuel moisture. *Global Change Biol.* **28**, 1544-1559 (2022).

10 Kelly, L. T. *et al.* Fire and biodiversity in the Anthropocene. *Science* **370**, eabb0355, doi:doi:10.1126/science.abb0355 (2020).

11 Filkov, A. I., Ngo, T., Matthews, S., Telfer, S. & Penman, T. D. Impact of Australia's catastrophic 2019/20 bushfire season on communities and environment. Retrospective analysis and current trends. *Journal of Safety Science and Resilience* **1**, 44-56, doi:https://doi.org/10.1016/j.jnlssr.2020.06.009 (2020).

12 Di Bitetti, M. S. Argentina: wildfires jeopardize rewilding. *Nature* **604**, 246-246, doi:10.1038/d41586-022-01006-5 (2022).

13 Senande-Rivera, M., Insua-Costa, D. & Miguez-Macho, G. Spatial and temporal expansion of global wildland fire activity in response to climate change. *Nature Communications* **13**, 1208, doi:10.1038/s41467-022-28835-2 (2022).

14 Alba, C., Skalova, H., McGregor, K. F., D'Antonio, C. & Pysek, P. Native and exotic plant species respond differently to wildfire and prescribed fire as revealed by meta-analysis. *Journal of Vegetation Science* **26**, 102-113, doi:10.1111/jvs.12212 (2015).
15 Jain, P. *et al.* A review of machine learning applications in wildfire science and management. *Environ. Rev.* **28**, 478-505, doi:10.1139/er-2020-0019 (2020).
16 Mahood, A. L., Koontz, M. J. & Balch, J. K. Fuel connectivity, burn severity, and seed bank survivorship drive ecosystem transformation in a semiarid shrubland. *Ecology* **104**, doi:10.1002/ecy.3968 (2023).
17 Pappalardo, P. *et al.* Comparing traditional and Bayesian approaches to ecological meta-analysis. *Methods Ecol. Evol.* **11**, 1286-1295, doi:https://doi.org/10.1111/2041-210X.13445 (2020).
18 Viechtbauer, W. Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software* **36**, 1-48 (2010).
19 Chard, M. *et al.* Time since fire influences macropod occurrence in a fire-prone coastal ecosystem. *Austral Ecol.* **47**, 507-518, doi:https://doi.org/10.1111/aec.13127 (2022).
20 Taillie, P. J. *et al.* Interacting and non-linear avian responses to mixed-severity wildfire and time since fire. *Ecosphere* **9**, e02291, doi:https://doi.org/10.1002/ecs2.2291 (2018).
21 Doherty, T. S., Macdonald, K. J., Nimmo, D. G., Santos, J. L. & Geary, W. L. Shifting fire regimes cause continent-wide transformation of threatened species habitat. *Proceedings of the National Academy of Sciences* **121**, e2316417121, doi:doi:10.1073/pnas.2316417121 (2024).
22 Steel, Z. L., Fogg, A. M., Burnett, R., Roberts, L. J. & Safford, H. D. When bigger isn't better—Implications of large high-severity wildfire patches for avian diversity and community composition. *Divers. Distrib.* **28**, 439-453, doi:https://doi.org/10.1111/ddi.13281 (2022).
23 Keith, D. A. *et al.* Fire-related threats and transformational change in Australian ecosystems. *Global Ecol. Biogeogr.* **31**, 2070-2084, doi:https://doi.org/10.1111/geb.13500 (2022).
24 Gosper, C. R. *et al.* Mapping risk to plant populations from short fire intervals via relationships between maturation period and environmental productivity. *Plant Ecol.* **223**, 769-787, doi:10.1007/s11258-022-01229-6 (2022).
25 Plumanns-Pouton, E. S., Swan, M. H., Penman, T. D., Collins, L. & Kelly, L. T. Time since fire shapes plant immaturity risk across fire severity classes. *Fire Ecology* **19**, doi:10.1186/s42408-023-00185-4 (2023).
26 Peterson, D. A. *et al.* Australia's Black Summer pyrocumulonimbus super outbreak reveals potential for increasingly extreme stratospheric smoke events. *npj Climate and Atmospheric Science* **4**, 38, doi:10.1038/s41612-021-00192-9 (2021).
27 Gibson, R. K. & Hislop, S. Signs of resilience in resprouting *Eucalyptus* forests, but areas of concern: 1 year of post-fire recovery from Australia's Black Summer of 2019–2020. *Int. J. Wildland Fire* **31**, 545-557, doi:https://doi.org/10.1071/WF21089 (2022).
28 von Takach, B. *et al.* Long-unburnt habitat is critical for the conservation of threatened vertebrates across Australia. *Landscape Ecol.* **37**, 1469-1482, doi:10.1007/s10980-022-01427-7 (2022).
29 Haslem, A. *et al.* Time-since-fire and inter-fire interval influence hollow availability for fauna in a fire-prone system. *Biol. Conserv.* **152**, 212-221, doi:10.1016/j.biocon.2012.04.007 (2012).
30 Robinson, N. M., Leonard, S. W. J., Bennett, A. F. & Clarke, M. F. Refuges for birds in fire-prone landscapes: The influence of fire severity and fire history on the distribution of forest birds. *For. Ecol. Manage.* **318**, 110-121, doi:https://doi.org/10.1016/j.foreco.2014.01.008 (2014).
31 Nimmo, D. G. *et al.* Predicting the century-long post-fire responses of reptiles. *Global Ecol. Biogeogr.* **21**, 1062-1073, doi:10.1111/j.1466-8238.2011.00747.x (2012).

32 Smith, A. L., Bull, C. M. & Driscoll, D. A. Successional specialization in a reptile community cautions against widespread planned burning and complete fire suppression. *J. Appl. Ecol.* **50**, 1178-1186, doi:10.1111/1365-2664.12119 (2013).

33 Miritis, V., Dickman, C. R., Nimmo, D. G. & Doherty, T. S. After the 'Black Summer' fires: Faunal responses to megafire depend on fire severity, proportional area burnt and vegetation type. *J. Appl. Ecol.* **61**, 63-75, doi:https://doi.org/10.1111/1365-2664.14545 (2024).

34 Lindenmayer, D. B. *et al.* Testing hypotheses associated with bird responses to wildfire. *Ecol. Appl.* **18**, 1967-1983 (2008).

35 Heard, G. W. *et al.* Drought, fire, and rainforest endemics: A case study of two threatened frogs impacted by Australia's "Black Summer". *Ecol. Evol.* **13**, e10069, doi:https://doi.org/10.1002/ece3.10069 (2023).

36 Marsh, J. R. *et al.* in *Australia's Megafires* (eds L. Rumpff *et al.*) 141-153 (CSIRO, 2023).

37 Moritz, M. A., Batllori, E. & Bolker, B. M. The role of fire in terrestrial vertebrate richness patterns. *Ecol. Lett.* **26**, 563-574, doi:https://doi.org/10.1111/ele.14177 (2023).

38 González, T. M., González-Trujillo, J. D., Muñoz, A. & Armenteras, D. Effects of fire history on animal communities: a systematic review. *Ecological Processes* **11**, 11, doi:10.1186/s13717-021-00357-7 (2022).

39 Penman, T. D., Clarke, H., Gibson, R. K., Collins, L. & Nolan, R. H. in *Australia's megafires, Biodiversity impacts and lessons from 2019-20* (eds L. Rumpff *et al.*) 42-59 (CSIRO, 2023).

40 Linley, G. D. *et al.* What do you mean, 'megafire'? *Global Ecol. Biogeogr.* **31**, 1906-1922, doi:https://doi.org/10.1111/geb.13499 (2022).

41 Ayars, J., Kramer, H. A. & Jones, G. M. The 2020 to 2021 California megafires and their impacts on wildlife habitat. *Proceedings of the National Academy of Sciences* **120**, e2312909120, doi:doi:10.1073/pnas.2312909120 (2023).

42 Takeshima, N. *et al.* Which is more generalizable, powerful and interpretable in meta-analyses, mean difference or standardized mean difference? *BMC Medical Research Methodology* **14**, 30, doi:10.1186/1471-2288-14-30 (2014).

43 Viechtbauer, W. *Metafor 4.7-11 Calculate Effect Sizes and Outcome Measures*, <https://wviechtb.github.io/metafor/reference/escalc.html> (2024).

44 user603. *(https://stats.stackexchange.com/users/603/user603), Detecting outliers in count data, URL (version: 2020-08-28): https://stats.stackexchange.com/q/56404*, 2020).

45 Morris, S. B. & DeShon, R. P. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods* **7**, 105-125, doi:10.1037/1082-989X.7.1.105 (2002).

46 Eales, J. *et al.* What is the effect of prescribed burning in temperate and boreal forest on biodiversity, beyond pyrophilous and saproxylic species? A systematic review. *Environmental Evidence* **7**, 19, doi:10.1186/s13750-018-0131-5 (2018).

47 Cuijpers, P., Weitz, E., Cristea, I. A. & Twisk, J. Pre-post effect sizes should be avoided in meta-analyses. *Epidemiol Psychiatr Sci* **26**, 364-368, doi:10.1017/s2045796016000809 (2017).

48 Collins, L. *et al.* The 2019/2020 mega-fires exposed Australian ecosystems to an unprecedented extent of high-severity fire. *Environmental Research Letters* **16**, 14, doi:10.1088/1748-9326/abeb9e (2021).

Reviewer Reports on the First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

The authors did a good job of responding to mine and other comments. My major suggestions are:

1. Improving the clarity of the figures.
2. Pumping the brakes on making inferences to every ecosystem on earth based on this study.
3. In my opinion, the cessation of fossil fuel combustion is orders of magnitude more important than any other management suggestion, and this study does indeed have the data behind it to at least say something along those lines.
4. Adding in diversity metrics (species richness, shannon-weiner etc) as response variables, if it is not too much trouble. (More of a minor suggestion)

Stated goal: "This extraordinary monitoring effort provided an unparalleled opportunity to quantify how extreme fire events affect biodiversity." Was this actually analyzed? I think a lot of people would be interested in the addition of simply looking at how the same sets of covariates affected things like species richness, species accumulation curves, alpha/beta diversity of the different taxonomic groups. Maybe that's easy to add, maybe it's a whole additional paper. Seems like one multipanel figure showing those basic things would be warranted.

The figures still have some room for improvement.

Suggestions for the Map figure: It is difficult to tell the difference between the different severities, and perhaps it is not all that necessary for severity to be shown in the map. Maybe simply polygonizing the fire pixels and showing the burned area extent would be easier to interpret, with an inset or two showing burn severity patterns.

Suggestions for the coefficient figures:
The division between high and low severity could be clarified even further with a legend, or simply "high severity" and 'low severity' on the side at a 90 degree angle.

But then there are other panels where the color is orange or red for the two severities, and then it just says high, mid low. So that is not immediately clear that it's high Pre-fire drought, for example. I think if all these figures are done the way the last one is, it would be better. e.g. Unburnt area: maybe just have e.g '> X ha' instead of "high", or just "High UA". Prefire drought: Maybe "PDSI >3'... etc. PA: '>70% PA'.

Maybe it would not be a bad idea to still split the non-interacting panels by severity, just to be consistent and less confusing. Just a thought.

It might be kind of interesting to have a histogram of the positive and negative effects, instead of just the numbers. It's hard to interpret if more sites/taxa etc are being negatively or positively affected. A lot of times it's kind of just half and half, but not always. Maybe a more efficient use of space would be in the row labels, you could have a ratio of negative, neutral and positive effects, e.g. (Alpine 55%/5%/30%)

Minor thought: Hard to squeeze all this data into a short article... Maybe consider a very long article as a companion to delve into the details more thoroughly.

General comment on the summary: In the context of the authors' response to my comment in the previous review about how there were no grasslands or shrublands in the study area, I think it is important to keep that in mind in the summary. e.g. line 124, "Building wildfire resilience into ecosystems depends on reducing fire occurrence". This is likely true in the ecosystems studied here, but not for all ecosystems, especially frequently burning grasslands or pine savannas. So I would try to build in that caveat in the abstract (although I understand the difficulty of adding nuance with a tight word limit). Maybe it's as simple as saying "these ecosystems" so it's clear you're not trying to say "all ecosystems". I also feel like "The fires had positive and negative impacts" is an important point to be made about the nuances of fire effects on different taxa (but also keeping in mind the short time after fire), but it is also vague, and could be more precisely worded to get at that nuance.

Line by line comments

126-130, future management, 461-465: given that pre-fire drought had overall negative effects, as did burn severity, and given that these two factors are driven at least partially by increasing temperatures due to climate change, maybe the cessation of fossil fuel consumption is an important thing to suggest? "Building wildfire resilience into ecosystems depends on reducing fire recurrence, which may require rapid wildfire suppression, especially in areas recently or frequently burnt. Defending wet ecosystems, expanding protected areas, and considering localized drought can also contribute." Is any of that really going to make much of a difference once we're hitting 2C and beyond?

199: Enright's 2015 interval squeeze paper could also be cited here.

210: it also implies that fire scientists need to be careful about saying 'prescribed fire is the answer to all of our problems'

287-288: This study does not have inference to all global ecosystems...

Referee #2 (Remarks to the Author):

The authors have done a good job of responding to most of my earlier comments/suggestions.

They have given deeper consideration to how their recommendations will be clashing or synergistic with other fire-prone parts of the world; this is evident in text, particularly at the end of the ms.

Please consider the following, with suggestions IN CAPS, for the abstract:

"...Building resilience into MANY ecosystems depends on reducing UNNATURALLY FREQUENT fire recurrence, which may require wildfire suppression, especially in areas recently or frequently burnt..."

(or something similar, to recognize/differentiate mixed conifer areas that are often assumed to have been dominated by frequent low severity fires)

Referee #3 (Remarks to the Author):

My comments have been addressed satisfactorily.

Author Rebuttals to First Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

The authors did a good job of responding to mine and other comments. My major suggestions are:

1. Improving the clarity of the figures.
2. Pumping the brakes on making inferences to every ecosystem on earth based on this study.
3. In my opinion, the cessation of fossil fuel combustion is orders of magnitude more important than any other management suggestion, and this study does indeed have the data behind it to at least say something along those lines.
4. Adding in diversity metrics (species richness, shannon-weiner etc) as response variables, if it is not too much trouble. (More of a minor suggestion)

Stated goal: "This extraordinary monitoring effort provided an unparalleled opportunity to quantify how extreme fire events affect biodiversity." Was this actually analyzed? I think a lot of people would be interested in the addition of simply looking at how the same sets of covariates affected things like species richness, species accumulation curves, alpha/beta diversity of the different taxonomic groups. Maybe that's easy to add, maybe it's a whole additional paper. Seems like one multipanel figure showing those basic things would be warranted.

Author Response:

Thank you for reviewing our manuscript again. In response to your suggestions:

1) We have revised our figures, please see our response to your below comment for details.
2) We have revised some of our text in response to your second comment, please see our detailed response below.
3) We have added in reference to climate change in response to your third comment, please see our detailed response below.
4) Further examination of diversity metrics is an interesting idea. We consider that biodiversity includes three levels of organisation; intra-specific/genetic variation, species, and interactions among species in communities and ecosystems [1]. We have examined patterns emerging from species-level responses, and discussed possible mechanisms involving species interactions representing the community level. We therefore feel that we have addressed biodiversity responses to wildfire, even though we did not examine community-level metrics, and that we are using the term 'biodiversity' appropriately.

While it will be very interesting to look at community level metrics such as richness, evenness and turnover, we think it is preferable not to expand the scope of this manuscript. It would necessarily increase the length of the manuscript, with additional explanation of methods and interpretation of any patterns that emerge and additional figures. However, our manuscript is already near the limit for word length and already has six figures or tables, so we cannot expand the scope.

We think these metrics could be more meaningfully addressed in a separate paper that had the space to fully explore a range of metrics, and to properly frame specific questions focussed on those metrics.

The figures still have some room for improvement.

Suggestions for the Map figure: It is difficult to tell the difference between the different severities, and perhaps it is not all that necessary for severity to be shown in the map. Maybe simply polygonizing the fire pixels and showing the burned area extent would be easier to interpret, with an inset or two showing burn severity patterns.

Author Response:
We agree that the colour gradient for fire severity was difficult to interpret in the original map. We have simplified the map by removing the colour gradient, with the map now showing fire extent.

Suggestions for the coefficient figures:
The division between high and low severity could be clarified even further with a legend, or simply "high severity" and 'low severity' on the side at a 90 degree angle.

Author Response:
The instructions to authors indicate "Where possible, text, including keys to symbols, should be provided in the legend rather than on the figure itself." Rather than add a legend, we have shifted the sentence describing what the orange and red row labels indicate to be the second sentence of the figure caption, making it easier for readers to find that critical information. We have also spelt out in full the high and low severity row labels in the top panel of the first results figures (Fig. 3), which provide readers with an initial guide to what the orange and red colours mean.

But then there are other panels where the color is orange or red for the two severities, and then it just says high, mid low. So that is not immediately clear that it's high Pre-fire drought, for example. I think if all these figures are done the way the last one is, it would be better. e.g. Unburnt area: maybe just have e.g '> X ha' instead of "high", or just "High UA". Prefire drought: Maybe "PDSI >3'… etc. PA: '>70% PA'.

Author Response:
We agree that adding more informative row labels will improve clarity. We have opted to use descriptive text, which aligns with our other response variables.

We have not used the alternative approach of including the category cutoffs, such as the suggested "PDSI >3". Those require longer labels in the second and third columns of the data figures, which would result in less room for the actual data. Also, such labels are not as easy to interpret. For example, for drought, the high drought values are less than -28.31, while low drought values are greater than -3.00. It would not be immediately obvious which is the high drought value if the label was "PDSI < - 28.31". We list the cut-off values between high, mid and low categories in Extended Data Table 1, and Extended Data Figure 6, which are now referred to in the caption for Fig. 3.

Maybe it would not be a bad idea to still split the non-interacting panels by severity, just to be consistent and less confusing. Just a thought.

Author Response:

We had these split in the original submission and this led to a difficult mismatch between the results presented in graphs, and the statistical test results. That is, we were presenting results for high and low severity in the graphs and mentioning these in the text, when the differences were not statistically supported. We therefore must retain the current presentation which only presents statistically supported relationships.

It might be kind of interesting to have a histogram of the positive and negative effects, instead of just the numbers. It's hard to interpret if more sites/taxa etc are being negatively or positively affected. A lot of times it's kind of just half and half, but not always. Maybe a more efficient use of space would be in the row labels, you could have a ratio of negative, neutral and positive effects, e.g. (Alpine 55%/5%/30%)

Author Response:
There are three reasons why we prefer not to add this information to the figures.

1. In our original submission our figures included three columns of sample size numbers, and this attracted substantial criticism for making the diagrams too complex. In view of that widely agreed initial problem, it seems best not to add more information into the figures.

2. The number or the ratio of responses contributing to the mean negative and mean positive effects is not a focus of our analysis and we do not mention those numbers in the text beyond identifying the sample sizes. The ratios may be interesting, but would require new analysis, more complex figures or new figures, and additional text to integrate into the manuscript.

3. The overall mean effects (the grey central column of graphs in Figs 3-5) provide a statistically rigorous summary of the distribution of effect sizes, both their size and their number. The suggested ratios of numbers of effects are therefore partially redundant.

Minor thought: Hard to squeeze all this data into a short article... Maybe consider a very long article as a companion to delve into the details more thoroughly.

Author Response:
We will consider additional articles, such as something with a focus on community metrics. However, we are happy with the length of the current article - the Nature format encouraging a streamlined presentation but within the context of the international literature. We feel the tight focus this format requires, has meant distilling the findings down to the most important points, which, in turn, makes our work more accessible to a broad audience.

General comment on the summary: In the context of the authors' response to my comment in the previous review about how there were no grasslands or shrublands in the study area, I think it is important to keep that in mind in the summary. e.g. line 124, "Building wildfire resilience into ecosystems depends on reducing fire occurrence". This is likely true in the ecosystems studied here, but not for all ecosystems, especially frequently burning grasslands or pine savannas. So I would try to build in that caveat in the abstract (although I understand the difficulty of adding nuance with a tight word limit). Maybe it's as simple as saying "these ecosystems" so it's clear you're not trying to say "all ecosystems". I also feel like "The fires had positive and negative impacts" is an important point to be made about the nuances of fire effects on different taxa (but also keeping in mind the short time after fire), but it is also vague, and could be more precisely worded to get at that nuance.

Author Response:
We have added 'these' as suggested, which clarifies our intention while also avoiding misinterpretation about other ecosystems. This elegant change also addresses the suggestion of Referee 2.

We have reworded the positive and negative impacts sentence in the summary which makes it easier for readers to understand.

Line by line comments

126-130, future management, 461-465: given that pre-fire drought had overall negative effects, as did burn severity, and given that these two factors are driven at least partially by increasing temperatures due to climate change, maybe the cessation of fossil fuel consumption is an important thing to suggest? "Building wildfire resilience into ecosystems depends on reducing fire recurrence, which may require rapid wildfire suppression, especially in areas recently or frequently burnt. Defending wet ecosystems, expanding protected areas, and considering localized drought can also contribute." Is any of that really going to make much of a difference once we're hitting 2C and beyond?

Author Response:
This is a good point. We have carefully revised the last part of the summary to make the key point about climate change, while staying within the 200 word limit.

"Building wildfire resilience into these ecosystems depends on reducing fire recurrence, including with rapid wildfire suppression in areas frequently burnt. Defending wet ecosystems, expanding protected areas, and considering localised drought could also contribute. While these countermeasures can help mitigate the impacts of more frequent megafires short-term, reversing anthropogenic climate change remains the urgent broad-scale solution."

We have also revised the closing paragraph of the main text to strengthen the message about the urgent need to address climate change.

"We stress however, that extreme fires have doubled over the past 20 years[2], coinciding with a 0.43 degree rise in the global temperature anomaly. Current global commitments to limit fossil fuel use will likely see a further 0.8 degree rise[3], with catastrophic worsening of extreme and extensive wildfires the predictable outcome. Urgent and massive reduction of fossil-fuel use remains an absolute priority."

199: Enright's 2015 interval squeeze paper could also be cited here.

Author Response:
We have replaced the two references we had used with the suggested Enright et al paper [4]. We have also edited the text to better represent what this paper says.

"Fire that is too frequent can make plants vulnerable to decline by reducing survival and limiting the time available to set seed[4]."

210: it also implies that fire scientists need to be careful about saying 'prescribed fire is the answer to all of our problems

Author Response:
This is true, and is a point we address in the Future Management section of the discussion.

287-288: This study does not have inference to all global ecosystems

Author Response:
We have reworded this to make the scope of inference match the data presented.

"If the global goal to protect 30% of land area by 2030 can be met in our study area, there is potential to increase resistance to wildfires, albeit slightly, contributing towards adapting to climate change."

Referee #2 (Remarks to the Author):

The authors have done a good job of responding to most of my earlier comments/suggestions.

They have given deeper consideration to how their recommendations will be clashing or synergistic with other fire-prone parts of the world; this is evident in text, particularly at the end of the ms.

Please consider the following, with suggestions IN CAPS, for the abstract:

"...Building resilience into MANY ecosystems depends on reducing UNNATURALLY FREQUENT fire recurrence, which may require wildfire suppression, especially in areas recently or frequently burnt..."

(or something similar, to recognize/differentiate mixed conifer areas that are often assumed to have been dominated by frequent low severity fires)

Author Response:
The addition of 'these ecosystems' as suggested by Referee 1 elegantly resolves this tension in the abstract.

Referee #3 (Remarks to the Author):

My comments have been addressed satisfactorily.

References

1 Redford, K. H. & Richter, B. D. Conservation of Biodiversity in a World of Use. *Conserv. Biol.* **13**, 1246-1256, doi:https://doi.org/10.1046/j.1523-1739.1999.97463.x (1999).

2 Cunningham, C. X., Williamson, G. J. & Bowman, D. M. J. S. Increasing frequency and intensity of the most extreme wildfires on Earth. *Nature Ecology & Evolution*, doi:10.1038/s41559-024-02452-2 (2024).

3 Meinshausen, M. *et al.* Realization of Paris Agreement pledges may limit warming just below 2 °C. *Nature* **604**, 304-309, doi:10.1038/s41586-022-04553-z (2022).

4 Enright, N. J., Fontaine, J. B., Bowman, D., Bradstock, R. A. & Williams, R. J. Interval squeeze: altered fire regimes and demographic responses interact to threaten woody species persistence as climate changes. *Front. Ecol. Environ.* **13**, 265-272, doi:10.1890/140231 (2015).

Reviewer Reports on the Second Revision:

Referees' comments:

Referee #1 (Remarks to the Author):

My comments have been satisfactorily addressed. Congratulations to the authors on a great contribution.