

Extensive climate-induced range shifts in butterflies across the globe

Corresponding Author: Dr Shawan Chowdhury

Version 0:

Decision Letter:

1st December 2025

*Please ensure you delete the link to your author homepage in this e-mail if you wish to forward it to your co-authors.

Dear Shawan,

Thank you for your patience while your Article entitled "Extensive climate-induced range shifts in butterflies across the globe" was under review. As you will see from the reports below, we have a split panel, as one of the reviewers is enthusiastic about the dataset and both have suggestions to improve the analysis, but they also both raise important technical concerns and one of the reviewers fails to see a strong advance for the field in the findings. After some discussion, the other editors and I agreed that we would like to see a revised version along with a response to the reviewers' criticisms.

We are committed to providing a fair and constructive peer-review process. Please do not hesitate to contact us if there are requests from the reviewers that you believe are technically impossible or unlikely to yield a meaningful outcome, while bearing in mind that we will be reluctant to consider the manuscript further in the absence of major revisions to make the analysis more rigorous and insightful. Adding to Reviewer 2's queries on the methods, please also provide a PRISMA flowchart (<https://www.prisma-statement.org/prisma-2020>).

If revising your manuscript:

* Please provide the manuscript in Microsoft Word format and highlight all changes in the same file or a copy.

* Include a "Response to reviewers" document detailing, point-by-point, how you addressed each referee comment. If no action was taken to address a point, you must provide a compelling argument. This response will be sent back to the referees along with the revised manuscript.

* If you have not done so already we suggest that you begin to revise your manuscript so that it conforms to our Article format instructions at <http://www.nature.com/natecolevol/info/final-submission>. Refer also to any guidelines provided in this letter.

* Include a revised version of any required reporting checklist. It will be available to referees (and, potentially, statisticians) to aid in their evaluation if the manuscript goes back for peer review. A revised checklist is essential for re-review of the paper.

We would hope to receive the revised manuscript within 6 months. If you cannot send it within this time, please let us know. We will be happy to consider your revision so long as nothing similar has been accepted for publication at Nature Ecology & Evolution or published elsewhere. Once ready, please use the link below:

Link Redacted

Note: This URL links to your confidential home page and associated information about manuscripts you may have submitted, or that you are reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage.

Nature Ecology & Evolution is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' on published papers create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS), prior to acceptance. This applies to primary research papers only. ORCID helps the scientific community achieve unambiguous

attribution of all scholarly contributions. You can create and link your ORCID from the home page of the MTS by clicking on 'Modify my Springer Nature account'. For more information please visit www.springernature.com/orcid.

Please do not hesitate to contact me if you have any questions or would like to discuss the required revisions further.

[redacted]

Reviewers' comments:

Reviewer #1 (Remarks to the Author):

The authors present a very important synthesis documenting widespread trends of range shifts in butterflies. The novelty in this work lies in the sheer taxonomic and geographic scope and efforts to incorporate non-English studies, all of which are impressive. I think this will be a valuable addition to the literature and will get cited a lot for those reasons. I also applaud the authors for expanding this study to look at "threats" in this group. There are a number of points below that I am concerned about before I believe this manuscript should proceed to publication; hopefully, the authors can address them.

Major Comments:

Why did the authors not extract information about the direction of shift? Particularly if much of the introduction motivates these range shifts by linking to climate change. What if a range shift was in the opposite direction that you would expect under warming conditions? In order to causally link these shifts, this feels important to me. Even if it's just documenting whether it is in the expected direction or not.

I find the section on the drivers of range shifts to be the most interesting but also unclear and somewhat problematic and unsatisfying for a couple of reasons. With regards to clarity, a minor comment is that it is unclear from the Supplementary Material whether the threats were the same as the reasons behind the range shift, or were they separate extractions? From the main text, it seems as though they are the same. If not, were the reasons categorical and what were the categories? Further, for the reasons behind the range shift, did the article have to establish a statistically significant relationship behind the reason? More broadly, I am concerned about the framing of the threats for two reasons. Firstly, the authors talk about threats having "the greatest impact" (L352-366) and "most dominant", but the effect sizes of these threats are not extracted from the studies nor compared to other threats. For example, many studies may have found a small effect of climate change, but fewer may have found an effect of habitat modification, but when the effect was present, its effect was much larger. Unless all studies are comparing all threats, can the authors really say that one threat has a greater impact than the others, versus that threat is just studied more often? This leads into my second point, that climate change may be more often linked to range shifts because it may be easier to quantify and model than other features like pollution and habitat loss. So just because it is more commonly cited as a driver does not mean it is the most important driver.

Another concern I have with this work is the potential effect that any publication bias would have on the findings: I would think it is much more likely that a study would get published (at least for a single species) if it documented some kind of range shift vs. null results (no shift). This would affect the denominator in many of your results (the total number of studies). (Also, is it reasonable to assume that expansions are just as likely to be published as contractions?) For this reason, it makes me hesitant to treat the findings about the proportion of butterflies shifting their ranges as any kind of absolute rates (including detection rates, e.g. L497). For example, stating that 80% of species have shifted their ranges may be an overestimate of the real-world proportions of species that actually have done so. Instead, I think it would be more defensible to treat the findings as relative rates: which geographic areas show proportionally more shifts, which family show proportionally more shifts, etc. Alternatively, if the authors could limit these statements to studies that looked at more than 1 species (so, more likely to be documenting community responses rather than focusing on single species, perhaps overcoming a lot of the publication bias), this could perhaps also be more robust.

Some of the wording throughout uses much more causal language that I think is appropriate. For example, L205 "range shift attribution" suggests the authors are aiming to do cause-and-effect statistical analyses, or test hypotheses about causality. L211: "Approximately 81% of species expanded their ranges horizontally primarily due to climate change and severe weather"; without knowing the details of all of these studies, it is perhaps safe to assume that many of these were correlative studies and perhaps did not compare all drivers (maybe only looked at climate change, but habitat loss could also explain some of the shift had it been included), so can the authors be confident that all of these shifts were causally driven by climate change, or were they linked to it?

Introduction: mentions these geographic biases and why they are a problem but it doesn't mention how the authors overcame this for their study.

L258: Is there a list of studies?

L299: "When considered relative to the species pool (i.e., the total number of butterfly species known to science, as reported in Pinkert et al.29), five European nations (i.e., Sweden, the Czech Republic, Finland, Luxembourg, and Spain) dominate,

each having records for >50% of their butterfly fauna" <- This statement is confusing to me because the first part of the sentence seems to sound like the global pool of all species, but then the second part refers to national species pools. Where did the national lists come from? I may have missed this somewhere, but I could not find it.

L298-314: I cannot tell what distinguishes these two approaches. In L298-306, it sounds like you are calculating the number of species with range shifts documented within that country divided by the total number of species in that country. In L307-314, are the authors also including species with range shifts documented within that country AND other countries in the numerator, and the same denominator?

L326: "The density of European records reflects both genuine changes and stronger detectability of range shifts given long-term biodiversity monitoring programs, whereas lower records from the Tropics likely reflect data scarcity rather than an absence of biotic responses" <- But shouldn't the PROPORTION of range shifts (out of all reports in the realm) be more important than the total numbers? This is of course assuming that the publication bias towards reporting a change is equal in all areas across the globe.

L455-470: I am definitely onboard with more citizen science efforts, but what I felt this section lacked was an explanation of why we need this AND also structured monitoring efforts (Step 1) – what would this provide that monitoring data could not?

L917: Is "bias" the right word here and throughout when referring to the preponderance of single-species studies? The fact that most studies only examine 1 species doesn't seem like a bias, just the current state of the field. Unless, when studies only look at one species, it may be more likely that is because they have detected a change, so single-species studies are less likely to report null results than community-wide studies that report trends including null shifts. Is it possible to summarize the trends separately for the single species studies vs non single species studies to see if the findings differ significantly?

Minor Comments:

It is unclear in the abstract and throughout the manuscript (e.g. L261-264) if range shift types were mutually exclusive, or a species could exhibit multiple types.

L236: A global analysis of 12415 species confirmed this bias <- unclear what "this bias" is; it sounds like you are talking about the geographical bias based on the previous sentence, but you switch to taxonomic bias here

L249: "their reliance on a narrow range of larval host plants and thermal tolerance," <- not all rely on a narrow range of host plants; many are quite generalist.. Also, the wording right now sounds like butterflies rely on thermal tolerance.

L260: "These reports 10% of the butterfly species" <- I think a word is missing here

L261: Didn't Pinkert et al. report 19,327 species?

L265: "with particularly high from the Tropics" <- missing a word here, or awkward sentence structure.

One thing that is unclear in the 'Global research effort in studying butterfly species range shifts section' is whether the pool of studies you are making statements about also includes studies that did investigate range shifts but documented no change (i.e. null results). For example: "However, there were substantially fewer reports of elevational range shifts in the Tropics": is that relative to other areas, and is that proportional to the number of studies that actually occurred there (i.e. effort)?

L315: Is this 80% the same result that is cited as 81% in the abstract?

L332-333: "In many tropical studies, range shifts were simplified as expansion or contraction when species newly appeared or disappeared locally" <- as opposed to what? I'm not sure what the alternative the authors are trying to compare to here.

L334-336: But also, the tropics are not warming as fast as temperate areas, so species don't have to move as far to maintain their climatic niche.

L338-340: unclear if this is specific to the tropics or a global estimate

Methods: why use both Web of Science and Google Scholar? Authors do not explain the motivation behind the dual effort. I had not heard of the term horizontal range shifts before, so its mention in the abstract confused me as I interpreted it as a longitudinal range shift. I think the use of the term horizontal range shifts in the main text is fine (after you define it), but perhaps consider using latitudinal and longitudinal range shifts in the abstract, or define it there as well.

Reviewer #2 (Remarks to the Author):

This study compiles butterfly range shift data, with a particular emphasis on under-studied regions like the tropics. They show substantial horizontal range expansions globally, with a prominence in tropical regions. They also assign key drivers of those range changes. I commend the authors for the amount of data they compiled, a much-needed effort. At the same time, I think the analysis and interpretation of results missed rigour and depth.

I thought the study would have been stronger had it been hypothesis-testing. There is lots of previous work to lean on. Related, the authors did not do a good enough job reviewing the literature. There was no background given for butterfly research specifically. There have been several big butterfly data papers they could have discussed. For example,

Hällfors MH, Heikkinen RK, Kuussaari M, Lehikoinen A, Luoto M, Pöyry J, Virkkala R, Saastamoinen M, Kujala H. Recent range shifts of moths, butterflies, and birds are driven by the breadth of their climatic niche. *Evol Lett.* 2023 Mar 12;8(1):89-100. doi: 10.1093/evlett/qrq004.

da Silva CRB, Diamond SE. Local climate change velocities and evolutionary history explain multidirectional range shifts in a North American butterfly assemblage. *J Anim Ecol.* 2024 Aug;93(8):1160-1171. doi: 10.1111/1365-2656.14132.

Antão, L. H., Weigel, B., Strona, G., Hällfors, M., Kaarlejärvi, E., Dallas, T., ... & Laine, A. L. (2022). Climate change reshuffles northern species within their niches. *Nature Climate Change*, 12(6), 587-592.

I also think they did not adequately describe variation in butterfly range shifts.

It's not clear why the authors used a vote-counting method instead of a meta-analysis. Vote-counting has many limitations. Magnitude of range expansions could vary a lot across species, quality of data.

What time frame is this data from? The time span can have many impacts.

Given the objective of including additional types of data, I found it surprising that the authors didn't look to see whether there was a difference in range shifts between English and non-English documents, or between expert assessments and from studies. There are more lessons to be learned from the new dataset.

Line 224- why are climate change impacts considered 'indirect'? There are many ways of thinking about 'indirect', please define

Line 230-231- Elaboration here needed

Line 240-241- Butterflies are not under-represented taxa

Line 326-329- Not clear what evidence there is that the European reflects "both genuine changes and stronger detectability"... "whereas lower records... reflect data scarcity rather than absence of biotic response." Why not compare the type of data between Tropics and Europe?

Results

Line 265-267- Was the diversity of the region taken into account with these estimates?

The results were not put into context with regional differences in temperature change.

What are implications of differences in taxonomic coverage?

Spatial patterns- What is the ecological relevance of the defining spatial patterns based on geopolitical boundaries?

Further, results are mixed and could be better explained. First describe, what % of country's regional pool have been assessed, then describe percentage of expansion, contraction.

Line 352-385- without "formal analyses" as the authors themselves acknowledge, I don't find the section on drivers of range shifts very compelling. How did authors evaluate the relative importance of threats?

I appreciate the aim of incorporating more types of data but more rigour is needed when presenting them as integrated results. The authors lean on single species examples.

Discussion

I found the discussion lacked substance and depth. Instead, it was a high-level summary of some mitigation strategies.

I didn't find the framework to be particularly novel or helpful, and it is disjunct from rest of paper. For example, no ideas were provided for how to actually integrate citizen science data and local knowledge.

Version 1:

Decision Letter:

31st March 2026

*Please ensure you delete the link to your author homepage in this e-mail if you wish to forward it to your co-authors.

Dear Shawan,

Your revised manuscript entitled "Extensive climate-induced range shifts in butterflies across the globe" has now been seen by the two reviewers. As you will see from the reports below, Reviewer 1 has a few remaining concerns to address with further revisions, while Reviewer 2 has no further requests. Therefore, we would like to see your responses to the points raised along with a revised manuscript before we can reach a final decision regarding publication.

When revising your manuscript:

* Please provide the manuscript in Microsoft Word format and highlight all changes in the same file or a copy.

* Include a "Response to reviewers" document detailing, point-by-point, how you addressed each reviewer comment. If no action was taken to address a point, you must provide a compelling argument. This response will be sent back to the reviewers along with the revised manuscript.

* If you have not done so already please begin to revise your manuscript so that it conforms to our Article format instructions at <http://www.nature.com/natecolevol/info/final-submission>. Refer also to any guidelines provided in this letter.

* Include a revised version of any required reporting checklist. It will be available to referees (and, potentially, statisticians) to aid in their evaluation if the manuscript goes back for peer review. A revised checklist is essential for re-review of the paper.

* Extended Data Figures - please ensure that any supplementary figures and tables that are crucial to the manuscript's conclusions are converted into Extended Data figures and tables to increase visibility of these data. Extended Data figures and tables are online-only (present in the online PDF and full-text HTML versions of the paper), peer-reviewed display items that provide essential background to the article but are not included in the main article due to space constraints. A maximum of ten Extended Data display items (figures and tables) is permitted.

Please use the link below to submit your revised manuscript and related files:

Link Redacted

Note: This URL links to your confidential home page and associated information about manuscripts you may have submitted, or that you are reviewing for us. If you wish to forward this email to co-authors, please delete the link to your homepage.

We hope to receive your revised manuscript within two months. If you cannot send it within this time, please let us know. We will be happy to consider your revision so long as nothing similar has been accepted for publication at Nature Ecology & Evolution or published elsewhere.

Nature Ecology & Evolution is committed to improving transparency in authorship. As part of our efforts in this direction, we are now requesting that all authors identified as 'corresponding author' on published papers create and link their Open Researcher and Contributor Identifier (ORCID) with their account on the Manuscript Tracking System (MTS), prior to acceptance. ORCID helps the scientific community achieve unambiguous attribution of all scholarly contributions. You can create and link your ORCID from the home page of the MTS by clicking on 'Modify my Springer Nature account'. For more information please visit <http://www.springernature.com/orcid>.

Please do not hesitate to contact me if you have any questions or would like to discuss these revisions further. I look forward to seeing the revised manuscript.

[redacted]

Reviewers' comments:

Reviewer #1 (Remarks to the Author):

I thank the reviewers for addressing most of my concerns in this revised version of the manuscript. I still have a number of smaller concerns that should be able to be addressed in another revision. Note that in the merged file sent to reviewers, there is no version of the manuscript with the changes highlighted or tracked, so it's difficult to assess where in the manuscript changes were made.

I think the authors have made it clear throughout most of the manuscript that any rates reported are not the true global prevalence of range shifts (and drivers) across all butterflies, and have instead framed them as the relative prevalence of documented shifts within their gathered evidence base. However, this has not carried into the abstract where I think it could be mentioned there. For much of it, it still reads as an evaluation of the true prevalence vs. the relative reporting rates.

Abstract: "Most species' (79%) range expansions were associated with climate change and severe weather" <- the placement of the 79% in the sentence: is that saying 79% of species expanded their ranges or of the range expansions documented, 79% were associated with climate change and severe weather?

Abstract: "We find stronger support for range expansions and contractions across tropical countries, with less evidence for elevational shifts" <- stronger compared to what? Compared to non-tropical or temperate regions?

L403-414: is this only horizontal contractions/expansions or poleward as well?

L347-350: Is there any reason to believe though that well-studied groups will show different patterns in shifts overall than lesser-studied taxa? Seems to me like it would only be a problem if one is not representative of the other, and the authors don't present any reason a priori that one taxonomic group might have different sensitivity/exposures compared to others.

L354-356: 80% + 19% <- what is the other 1%?

L363: Why did the authors need to extract the country's centroid if they were just grouping countries into Europe, the tropics, and non-European, non-tropical countries?

L389-392: "The density of European records reflects both genuine changes and stronger detectability of range shifts given long term biodiversity monitoring programs, whereas the lower density of range-shift records from the tropics likely reflects data scarcity rather than an absence of biotic responses." <- I find this a bit of a reach; the authors cannot separate the detection vs. true prevalence of range shifts in your data, so what is the evidence behind this statement?

L397: "An important note is that many tropical studies, range-shifts were simplified as expansion or contraction evidence when species newly appeared or disappeared locally (i.e., by comparing checklists rather than relying on quantitative data and proper statistical tests; Table 1), obscuring patterns in these regions." <- I think there is a word missing in this sentence, between "that" and "many"; perhaps "in"? Secondly, I could see how that would be an overall problem, but it doesn't obscure your own patterns, correct? Because you did not extract any of the quantitative data from these studies (just the type of range shift). So this statement seems oddly phrased or placed, because the previous sentence is about your specific study.

L432: "The relative importance of different threat categories varied across range-shift types" <- I am not sure that the words "relative importance" are appropriate because (as the authors themselves state) the study cannot investigate the magnitude of impact or causal dominance of each of the threats. Suggest rewording to something like "the relative reporting rates".

I wonder if for Figure 4B, this should be a relative proportion rather than absolute numbers. I worry that absolute numbers obscure: a) the difference in the sizes of species' pool (i.e. more species in the tropics); and b) the number of assessments per continent. Low numbers of documented threats like in Africa and South America may just reflect a lower number of studies that examined threats rather than actual prevalence of threats.

L432-436: What was the most frequently reported threat for species with documented range expansions?

L662-664: "an information that was either reported by a given study or that was assessed directly by our set of butterfly experts who relied on field observations" <- I think there is a typo in this sentence around "an information"

L659-667: "Once each expert agreed, we asked them to share the same type of range shift information and to fill the same form by providing the type of range shift (i.e., distinguishing between horizontal range expansion, horizontal range contraction and elevational range shifts) (see Table 1), an information that was either reported by a given study or that was assessed directly by our set of butterfly experts who relied on field observations. Specifically, the experts compared the area of occupancy maps covered by field observations (see Table 1) to determine whether there has been a change in the species range and which type of range changes among the three studied range-shift types." <- I am confused about the expert elicitation process throughout the manuscript. When studies were provided to an expert, how were they matched to an expert? Why did they need an expert to judge rather than one of the regular data extractors. When it was "assessed directly by our set of butterfly experts who relied on field observations", was observation data provided TO the experts or BY the experts? And they based their judgment solely off of that, and no prior knowledge of the species? How were studies or species selected for each expert? Finally, the second mention of "see Table 1" seems misplaced as it has nothing to do with that section of the sentence.

L676: Reference #30 is not a taxonomy; is this the right citation here?

L676-677: From this wording, it seems like range shift information was extracted from the original study and this reference was used to extract country and country-level distribution data. I don't understand why country wasn't extracted from the original study/observation itself, and why it needed to be done in a posthoc manner like this. What if a species spanned multiple countries? Was a range shift counted in each even if the original study only documented it in one country?

There is no Supplementary Figure S1 (cited in the Supplementary Material)in any of the documents provided to reviewers.

Reviewer #2 (Remarks to the Author):

I appreciate the extent of revisions made by the authors. I am content with the revisions made and am happy to recommend the manuscript be published.

Version 2:

Decision Letter:

21st April 2026

Dear Shawan,

Thank you for submitting your revised manuscript "Extensive climate-induced range shifts in butterflies across the globe" (NATECOLEVOL-25103736B). Based on the feedback from Reviewer 1 (see their report below), we will be happy in principle to publish it in Nature Ecology & Evolution, pending minor revisions to satisfy the reviewers' final requests and to comply with our editorial and formatting guidelines.

We are now performing detailed checks on your paper and will send you a checklist detailing our editorial and formatting requirements in about a week. Please do not upload the final materials and make any revisions until you receive this additional information from us.

[redacted]

Reviewer #1 (Remarks to the Author):

I thank the authors for their continued diligence in addressing my concerns. Just a few additional minor comments at this point:

L210-214: "Most species (80%) experienced range expansions, and most range shifts (79%) were associated with climate change and extreme weather events. A substantial proportion of species in our dataset contracted their ranges (27%) or shifted along elevational gradients (22%)." <- are all of these numbers reported in the main text? I could not find the 27%, 22% or 80% figures in the main text.

L276: "Our dataset comprises 6182 range shift reports" <- does this number include instances of negative findings (i.e. species who did not experience a range shift)? If so, would suggest using a different term here than "range shift reports" as it may imply only positive instances. Similar feedback for L280-282. Do these numbers include instances of absences of horizontal range expansions, etc. or are they only the number of records with positive evidence?

L312-313: "Relative to the total number of known species per family, nymphalids had the highest percentage of species shifting their ranges, while riodinids had the lowest" <- wouldn't it also make sense to make it relative to the number of reports in your dataset, which would alleviate concerns about uneven effort between taxonomic groups?

L313-317: I appreciate the authors' clarification here. In their response, they address how uneven effort can affect detection and reporting of range shifts between groups. I would note that having uneven sampling effort between families would only underestimate the prevalence of range shifts in the understudied family if the denominator was the total number of species in a family, rather than the total number of reports in a family (see above comment). However, it is the word "bias" that still stands out to me here in this section, as well as the link to the actual drivers of range shifts (not just prevalence): the authors claim that "uneven taxonomic coverage can bias global inference about the drivers and consequences of species range shifts". However, this would only introduce a bias if one family is not representative of the other, and the authors don't present any reason a priori that one taxonomic group might have different sensitivity/exposures compared to others, no a priori reason that this would bias drivers.

L373-375: "Because several large countries (e.g., USA, Australia) span both tropical and temperate zones: as an example here, USA was allocated to the non-European countries group." <- the sentence structure here is awkward and grammatically incorrect.

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

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

The images or other third party material in this Peer Review File are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

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

Revisions for “Extensive climate-induced range shifts in butterflies across the globe”

We thank the Editor and two anonymous reviewers for their helpful comments. We have revised the manuscript accordingly. Below, we detail our responses and changes to the manuscript (MainManuscript_revised_trackChanged; in blue).

REVIEWER 1

The authors present a very important synthesis documenting widespread trends of range shifts in butterflies. The novelty in this work lies in the sheer taxonomic and geographic scope and efforts to incorporate non-English studies, all of which are impressive. I think this will be a valuable addition to the literature and will get cited a lot for those reasons. I also applaud the authors for expanding this study to look at “threats” in this group.

Response: We thank the Reviewer for their very positive assessments and their encouraging comments. This is highly appreciated. By highlighting the added value of reports from non-English studies and sources, including experts’ knowledge, we sincerely hope that it will foster further initiatives to retrieve various sources of data and knowledge that still remain untapped so far.

Major Comments:

Why did the authors not extract information about the direction of shift? Particularly if much of the introduction motivates these range shifts by linking to climate change. What if a range shift was in the opposite direction that you would expect under warming conditions? In order to causally link these shifts, this feels important to me. Even if it’s just documenting whether it is in the expected direction or not.

I find the section on the drivers of range shifts to be the most interesting but also unclear and somewhat problematic and unsatisfying for a couple of reasons. With regards to clarity, a minor comment is that it is unclear from the Supplementary Material whether the threats were the same as the reasons behind the range shift, or were they separate extractions? From the main text, it seems as though they are the same. If not, were the reasons categorical and what were the categories?

Further, for the reasons behind the range shift, did the article have to establish a statistically significant relationship behind the reason? More broadly, I am concerned about the framing of the threats for two reasons. Firstly, the authors talk about threats having “the greatest impact” (L352-366) and “most dominant”, but the effect sizes of these threats are not extracted from the studies nor compared to other threats. For example, many studies may have found a small effect of climate change, but fewer may have found an effect of habitat modification, but when the effect was present, its effect was much larger. Unless all studies are comparing all threats, can the authors really say that one threat has a greater impact than the others, versus that threat is just studied more often? This leads into my second point, that climate change may be more often linked to range shifts because it may be

easier to quantify and model than other features like pollution and habitat loss. So just because it is more commonly cited as a driver does not mean it is the most important driver.

Response: We thank the Reviewer for raising these valuable points. We agree that directionality (e.g., poleward vs equatorward in latitude or eastward vs westward in longitude; upslope vs downslope) is informative when assessing whether documented range shifts align with expectations under global warming. However, we were not able to extract range shift directions because this information was not reported consistently across the heterogeneous evidence base, particularly in non-English sources and expert assessments, where range changes are often documented as local appearances/disappearances or checklist updates without spatially explicit information on range positions or direction. In addition, our synthesis operates primarily at the country level rather than from paired historical–contemporary locations or surveys, so any directional change classification would be coarse and potentially misleading, as local responses may differ within countries. We have therefore revised the wording throughout, added explanations on the actual data at hand (i.e., the lack of precise and quantitative information on range shift rates and directions which would only be available for a very small subset of countries in our assembled dataset), and added sensitivity analyses to better align our inferences with the data at hand.

Regarding how the information on drivers of range shifts was compiled, we extracted the textual reasons/attributions reported in each source for the observed range shift, which were then mapped to the IUCN standardised threat categories.

For the sake of clarity and transparency, we have moved the following paragraph from the Supplementary Methods to the main text [Lines 707–713].

“From each study, we extracted the publication information (e.g., journal, DOI), taxonomic details (e.g., species), geographic information (e.g., country), type of range shift (e.g., expansion/contraction), the potential drivers behind the range shift type (as stated by the authors and experts), whether the species is invasive, the impact on native species, and whether the species is facing any threat. We used countries because most sources report shifts at the national level (atlases/checklists/expert assessments) and because national baselines are required for conservation reporting and resource allocation.”

We agree that our original wording (‘greatest impact’, ‘most dominant’) could be interpreted as ranking drivers by effect or magnitude, whereas our synthesis was intended to summarise the prevalence of reported driver attributions across studies. We have therefore revised the manuscript accordingly to replace ‘impact/dominance’ language with ‘most frequently reported’ / ‘reported to have the highest impact’ [e.g., Line 466, 477]. We have also added an explicit clarification that our driver/threat synthesis does not compare effect sizes or causal dominance [Lines 432–434], because most studies did not estimate commensurate multi-driver effect sizes and many did not evaluate multiple drivers simultaneously. We have added the following recommendation on lines 515–518.

“Future work could combine standardised climate metrics (e.g., regional warming rates or climate-change velocity) and land-cover metrics to estimate the magnitude of range shifts and test whether the direction and rate of change align with regional exposure to environmental change.”

Finally, we also agree that some drivers (particularly climate change) may be reported more frequently than others, partly because they are easier to quantify and model. To address this important comment, we have now conducted an effort-adjusted sensitivity analysis that downweights references reporting fewer threat categories relative to those that report and discuss several threat categories. In the revised version of our manuscript, we now present both species-weighted (reference $\times$ species $\times$ threat) and reference-weighted (reference $\times$ threat) summaries to assess whether our results could also be driven by species-rich references. We have added the following paragraph in the methods [Lines 736-744].

“Because most studies did not estimate comparable multi-driver effect sizes (and many did not evaluate multiple drivers simultaneously), we summarised drivers as the prevalence of reported attributions rather than impact magnitude. To assess whether patterns could reflect differences in the breadth of reported drivers, we performed an effort-adjusted sensitivity analysis in which each study’s driver attributions were weighted by the inverse of the number of distinct threat categories reported in that study. We reported results (Extended Data 7) using (i) a species-weighted summary (reference $\times$ species $\times$ threat) and (ii) a reference-weighted summary (reference $\times$ threat), to evaluate the influence of references with a higher number of reports from multiple species.”

Our results were qualitatively robust across these different summaries and sensitivity analyses (Extended Data 7), indicating that our main conclusions about the relative prevalence of reported drivers are not solely an artefact of single-driver reporting. We have added the following paragraph in lines 494-504.

“These patterns may reflect both ecological differences and differential measurability and study emphasis among drivers (e.g., climate covariates are often more readily available than local pollution or fine-scale habitat degradation). However, an effort-adjusted sensitivity analysis produced qualitatively similar driver patterns to those based on raw counts, for both species-weighted and reference-weighted summaries (Extended Data 7). Where rankings differed between summaries, the differences reflected whether a threat was concentrated in a small number of species-rich references or distributed broadly across references, rather than evidence of comparative effect magnitude. Overall, climate change and severe weather is the most frequently reported threat category linked to butterfly range shifts in our database, and this ranking is robust to normalising for differences in the number of threat categories reported per reference.”

Another concern I have with this work is the potential effect that any publication bias would have on the findings: I would think it is much more likely that a study would get published (at least for a single species) if it documented some kind of range shift vs. null results (no shift). This would affect the denominator in many of your results (the total number of studies). (Also, is it reasonable to assume that expansions are just as likely to be published as contractions?) For this reason, it makes me hesitant to treat the findings about the proportion of butterflies shifting their ranges as any kind of absolute rates (including detection rates, e.g. L497) . For example, stating that 80% of species have shifted their ranges may be an overestimate of the real-world proportions of species that actually have done so. Instead, I think it would be more defensible to treat the findings as relative rates:

which geographic areas show proportionally more shifts, which family show proportionally more shifts, etc. Alternatively, if the authors could limit these statements to studies that looked at more than 1 species (so, more likely to be documenting community responses rather than focusing on single species, perhaps overcoming a lot of the publication bias), this could perhaps also be more robust.

Response: See our detailed response to this comment just below, after the next paragraph written by reviewer #1, which is about the same concern.

L917: Is “bias” the right word here and throughout when referring to the preponderance of single-species studies? The fact that most studies only examine 1 species doesn’t seem like a bias, just the current state of the field. Unless, when studies only look at one species, it may be more likely that is because they have detected a change, so single-species studies are less likely to report null results than community-wide studies that report trends including null shifts. Is it possible to summarize the trends separately for the single species studies vs non single species studies to see if the findings differ significantly?

Response: We thank the Reviewer for raising this important issue of publication and reporting bias and for providing constructive comments to help us address this issue. We agree that studies documenting a clear range shift—particularly single-species studies—may be more likely to be published than studies reporting null or negative results. We have therefore:

a) revised the manuscript to avoid presenting these proportion values as detection rates or global prevalence across all butterflies, and instead framed them as the prevalence of documented shifts within our gathered evidence base for comparative purposes among countries. For example, we have added the following sentences:

“Importantly, these percentages quantify documentation coverage in our evidence database, not the absolute prevalence of range shift records across all butterflies within a country.” [Lines 376-378]

“We then examined the composition of documented range-shift types to avoid conflating differences in the volume of evidence with differences in range shifts.” [Lines 389-390]

b) conducted a sensitivity analysis that (i) summarises patterns at the study level (with each reference counted once per shift type, reducing the influence of species-rich studies) and (ii) restricts the dataset to multi-species references (≥ 2 species per reference), which are less likely to be motivated by a single species-level positive result. We have added the following paragraph in lines 415-427.

“To assess sensitivity to a potential publication bias (e.g., preferential publication of striking single-species shifts), we repeated the continental summaries using a reference-weighted approach (each reference counted once per range-shift type) and restricted the analysis to multi-species references (≥ 2 species per reference). Patterns were broadly consistent, with horizontal range expansions remaining the most commonly reported shift type across continents. However, in the multi-species subset, the relative contribution of horizontal range contractions increased in some regions (e.g., expansions dropped from 0.93 to 0.73 in Africa, and contractions increased from 0.3 to 0.36 in Europe; Extended Data 6). These

results support interpreting our findings as relative patterns within the published record rather than as absolute rates of change. Future studies could conduct a targeted meta-analysis of range-shift magnitude for the subset of studies reporting compatible quantitative outcomes (e.g., standardised range shift rates and uncertainties, which we were not able to compile)."

We have deleted the 'bias' term from Lines 1171-1173.

Some of the wording throughout uses much more causal language that I think is appropriate. For example, L205 "range shift attribution" suggests the authors are aiming to do cause-and-effect statistical analyses, or test hypotheses about causality. L211: "Approximately 81% of species expanded their ranges horizontally primarily due to climate change and severe weather"; without knowing the details of all of these studies, it is perhaps safe to assume that many of these were correlative studies and perhaps did not compare all drivers (maybe only looked at climate change, but habitat loss could also explain some of the shift had it been included), so can the authors be confident that all of these shifts were causally driven by climate change, or were they linked to it?

Response: We agree that some phrasing in the manuscript could be interpreted as implying causal inference. We have therefore substantially revised the abstract and the manuscript to remove causal language [e.g., lines 215-238], and as discussed in the previous response, we have substantially revised the 'drivers' section to address the reviewer's comments. Our study is a global synthesis of multiple types of evidence rather than a meta-analysis, and we did not evaluate causal mechanisms or compare effect sizes across drivers.

Introduction: mentions these geographic biases and why they are a problem but it doesn't mention how the authors overcame this for their study.

Response: We have added the following information on lines 267-272.

"To address these limitations, we focused on butterflies—the very first group used as a sentinel of change to detect comprehensive range-shifts signals in response to global warming³. We combined systematic literature searches with targeted extraction of evidence from non-English sources, grey literature, and expert knowledge, thereby expanding distribution data coverage beyond regions and taxa that dominate conventional literature reviews."

L258: Is there a list of studies?

Response: We have added the list to the Extended Data 2.

L299: “When considered relative to the species pool (i.e., the total number of butterfly species known to science, as reported in Pinkert et al.29), five European nations (i.e., Sweden, the Czech Republic, Finland, Luxembourg, and Spain) dominate, each having records for >50% of their butterfly fauna” <- This statement is confusing to me because the first part of the sentence seems to sound like the global pool of all species, but then the second part refers to national species pools. Where did the national lists come from? I may have missed this somewhere, but I could not find it.

Response: We apologise for the confusion. Here, the denominator is the national butterfly species pool for each country (i.e., the number of butterfly species recorded from that country), and the numerator is the number of those national species for which we found at least one documented range-shift record in our synthesis. We used country-level species richness/checklist totals compiled by Pinkert et al. to represent the national species pools. We have added ‘country-level’ in line 367 for clarity.

L298-314: I cannot tell what distinguishes these two approaches. In L298-306, it sounds like you are calculating the number of species with range shifts documented within that country divided by the total number of species in that country. In L307-314, are the authors also including species with range shifts documented within that country AND other countries in the numerator, and the same denominator?

Response: Once again, we apologise for the confusion. Suppose a given country (X) has 100 species of butterflies as the total species pool. We obtained documented range shifts for 50 species from X. According to Fig. 2, 50% of butterfly species in that country have experienced some form of range shift.

However, the observed range shift in country X could also be due to the fact that other species from the species pool in country X were never studied in country X, but elsewhere in other countries from which we have range shift reports. Let us assume that country X has 20 additional species that have been recorded as experiencing range shifts in another country (Y), but not in X. According to Fig. 3, 70% of butterfly species (documented 50 + undocumented 20) in that country have experienced some form of range shift, either reported directly in country X or elsewhere (i.e., country Y).

For clarity, we have revised the first part of the paragraph, and it now reads as:

“To better address sampling gaps in countries, we calculated the percentage of butterfly species per country for which we have range shift data, but this time relative to the set of species that have at least one documented range shift record anywhere in our dataset (i.e., species with at least one record of expansion, contraction, or elevational shift in any country)” [Lines 379-384]

L326: “The density of European records reflects both genuine changes and stronger detectability of range shifts given long-term biodiversity monitoring programs, whereas

lower records from the Tropics likely reflect data scarcity rather than an absence of biotic responses” <- But shouldn't the PROPORTION of range shifts (out of all reports in the realm) be more important than the total numbers? This is of course assuming that the publication bias towards reporting a change is equal in all areas across the globe.

Response: We agree that absolute record density is partly an effort signal and can be misleading when interpreted as a measure of biological prevalence. Our intention was to describe the concentration of evidence, which is influenced by long-term monitoring capacity and the availability of sources in each country. We have now clarified the wording and, as suggested, we now also report proportion-based summaries (normalised within continent/realm) so that differences in total reporting volume do not bias interpretations (Extended Data 3).

In addition, to address the suggestion of reviewer 2 (see below), we performed a region-stratified comparison of the documented number of range-shift species across Europe, tropical countries, and non-European non-tropical countries, further partitioned by range shift type (horizontal expansion, horizontal contraction, elevational shifts). We have added the following paragraph to lines 403-414 and a figure to Extended Data 5.

“To quantify whether the reports of range-shifts in our database differed across regions, we extracted the country centroid, grouped countries into Europe, the tropics, and non-European, non-tropical countries. Shift-type composition differed significantly among regions (Pearson’s $\chi^2 = 389.13$, $df = 4$, $p = < 0.001$; Cramér’s $V = 0.277$), driven by an over-representation of expansions in the tropics and higher relative representation of contractions and elevational shifts in Europe. In Europe, documented records was relatively balanced across expansions (41.3%), contractions (37.4%), and elevational shifts (21.3%). By contrast, expansions dominated in the tropics (80.3%) and in non-European non-tropical countries (64.8%) (Extended Data 5). Importantly, elevational range shifts were rarely documented in tropical countries, highlighting a major gap in the current evidence base and reinforcing the need for integrated monitoring approaches that improve detection of under-reported range-shift types.”

L455-470: I am definitely onboard with more citizen science efforts, but what I felt this section lacked was an explanation of why we need this AND also structured monitoring efforts (Step 1) – what would this provide that monitoring data could not?

Response: We thank the reviewer for their thoughtful discussion. We have added the following information in lines 631-637.

“However, these data are often opportunistic and heterogeneous in effort, making them difficult to use alone for robust inference about temporal trends, the rate of range change, or the mechanisms underpinning species range shifts. For this, citizen science is most powerful when paired with structured monitoring (Step 1): standardised protocols provide consistent effort through time and space, improving range-shift detectability, and generating comparable baselines against which opportunistic observations can be calibrated¹¹⁴.”

We have also added an example from our recent study, highlighting the potential of citizen science data.

“For instance, using the *Acraea terpsicore* butterfly, a recent study found that integrating species locality records from social media with records from the biodiversity database can track expansion much more effectively, identifying new distributions from the tropics, especially from high-elevation areas¹¹³.” [Lines 627-631]

Minor Comments:

It is unclear in the abstract and throughout the manuscript (e.g. L261-264) if range shift types were mutually exclusive, or a species could exhibit multiple types.

Response: Each species could be reported to have different types of shifts, as mentioned, for example, in lines 394-399.

L236: A global analysis of 12415 species confirmed this bias <- unclear what “this bias” is; it sounds like you are talking about the geographical bias based on the previous sentence, but you switch to taxonomic bias here

Response: We have changed ‘confirmed the bias’ to ‘identified’ for clarity [Line 258].

L249: “their reliance on a narrow range of larval host plants and thermal tolerance,” <- not all rely on a narrow range of host plants; many are quite generalist.. Also, the wording right now sounds like butterflies rely on thermal tolerance.

Response: We have rephrased the sentence. It now reads as:

“Butterflies (~19000 species, globally) occupy diverse zoogeographic regions, altitudes and habitats²⁸. Their well-characterised larval host-plant associations, high sensitivity to temperature, relatively high dispersal capacity, and short lifecycles, together render them a highly effective early warning system for detecting the biological impacts of climate change^{29,30,31,32}.” [Lines 258-264]

L260: “These reports 10% of the butterfly species” <- I think a word is missing here

Response: We have added ‘documented’ in line 305.

L261: Didn't Pinkert et al. report 19,327 species?

Response: We thank the reviewer for pointing this out. We have revised this in line 305.

L265: "with particularly high from the Tropics" <- missing a word here, or awkward sentence structure.

Response: We have added 'records' in line 310.

One thing that is unclear in the 'Global research effort in studying butterfly species range shifts section' is whether the pool of studies you are making statements about also includes studies that did investigate range shifts but documented no change (i.e. null results). For example: "However, there were substantially fewer reports of elevational range shifts in the Tropics": is that relative to other areas, and is that proportional to the number of studies that actually occurred there (i.e. effort)?

Response: Our evidence base is compiled from sources that reported at least one butterfly range shift event (expansion, contraction, or elevational shift) and therefore does not systematically include studies that investigated range shifts but concluded no detectable change (null results). We have added the following limitation in lines 376-378.

"Importantly, these percentages quantify documentation coverage in our evidence database, not the absolute prevalence of range shift records across all butterflies within a country."

L315: Is this 80% the same result that is cited as 81% in the abstract?

Response: No, it's not the same. The abstract is talking about "*Climate change and severe weather* was linked to range expansion for 80% of species, whereas this part is talking about "Most species (80%) had records for a single type of range shift in our dataset" [Line 393]."

L332-333: "In many tropical studies, range shifts were simplified as expansion or contraction when species newly appeared or disappeared locally" <- as opposed to what? I'm not sure what the alternative the authors are trying to compare to here.

Response: For clarity, we have added the following information in lines 439-440: "by comparing checklists rather than relying on quantitative data and proper statistical tests".

L334-336: But also, the tropics are not warming as fast as temperate areas, so species don't have to move as far to maintain their climatic niche.

Response: We have added the following sentence in lines 443-445.

"However, the tropics are not warming as fast as temperate areas, which could limit the number of range-shift reports of butterfly species from the tropics."

L338-340: unclear if this is specific to the tropics or a global estimate

Response: We have clarified that this is a global study in line 447.

Methods: why use both Web of Science and Google Scholar? Authors do not explain the motivation behind the dual effort.

Response: We have revised the following statements for clarity in lines 698-704.

"Many authors publish their research in regional journals and in non-English languages and these studies are vastly unavailable in the international search systems, such as Web of Science^{42,131}. For this, we searched for published literature in 15 languages (see Extended Data 1; Extended Data 8). For English-language studies, we used Web of Science and Google Scholar, while for studies published in non-English languages, we used Google Scholar and local search systems available within the focal country (see supplementary methods)."

I had not heard of the term horizontal range shifts before, so its mention in the abstract confused me as I interpreted it as a longitudinal range shift. I think the use of the term horizontal range shifts in the main text is fine (after you define it), but perhaps consider using latitudinal and longitudinal range shifts in the abstract, or define it there as well.

Response: We have deleted 'horizontal' from the abstract [line 223].

REVIEWER 2

This study compiles butterfly range shift data, with a particular emphasis on under-studied regions like the tropics. They show substantial horizontal range expansions globally, with a prominence in tropical regions. They also assign key drivers of those range changes. I commend the authors for the amount of data they compiled, a much-needed effort.

Response: We thank the Reviewer for their very positive assessments and their encouraging comments. This is highly appreciated.

At the same time, I think the analysis and interpretation of results missed rigour and depth. I thought the study would have been stronger had it been hypothesis-testing. There is lots of previous work to lean on. It's not clear why the authors used a vote-counting method instead of a meta-analysis. Vote-counting has many limitations. Magnitude of range expansions could vary a lot across species, quality of data.

Response: We thank the Reviewer for this important feedback. We fully agree that a meta-analysis framework would be a more powerful approach than the semi-quantitative approach we undertook. Yet, implementing a meta-analysis would require that estimates of range shifts and their uncertainty be provided in the original studies we retrieved from the scientific and grey literature, as well as the experts' knowledge we assembled. However, such quantitative information is simply missing from the vast majority of the sources we gathered here. Our initial goal was to conduct a global evidence synthesis of butterfly range shift records and drivers across diverse evidence types, such as peer-reviewed papers, faunal atlases/books, grey literature, and expert assessments, and in multiple languages. Even when some quantitative information was reported, it was rarely standardised, and methods were also not harmonised (e.g., differing time windows, spatial units, detection/occupancy frameworks, and definitions of range edge). Consequently, implementing a formal meta-analysis would require restricting our current analyses to a much smaller, non-representative subset of studies and would likely replicate geographic and taxonomic biases in other studies. Given the lack of systematic monitoring data on range shifts and the known biases that neglect non-vertebrate taxa and Global South regions, our motivation was to gather as much evidence as possible on butterfly redistribution, including from different types of evidence and even qualitative statements, aiming to get as broad a picture of this phenomenon as possible. While we believe this effort has provided robust evidence about the patterns of range shifts, it does not allow for accurate quantification of the magnitude or the velocity at which butterfly species have been changing their distributions. We note that drawing on multiple types of evidence is a powerful approach to quantify biodiversity change for underrepresented taxa and /or regions (Cooke et al. 2025, Clements et al. 2025).

We have added several sentences in the revised manuscript (e.g., lines 376-378), to better explain our initial motivations and be clear on what we can, and cannot, achieve with the data we have gathered, while leaving the door open for further initiatives to dig deeper and improve our knowledge on the rate and direction of butterfly species range shifts.

“Future work could combine standardised climate metrics (e.g., regional warming rates or climate-change velocity) and land-cover metrics to estimate the magnitude of range shifts and test whether the direction and rate of change align with regional exposure to environmental change.” [Lines 515-518]

Related, the authors did not do a good enough job reviewing the literature. There was no background given for butterfly research specifically. I also think they did not adequately describe variation in butterfly range shifts. There have been several big butterfly data papers they could have discussed. For example,

Hällfors MH, Heikkinen RK, Kuussaari M, Lehikoinen A, Luoto M, Pöyry J, Virkkala R, Saastamoinen M, Kujala H. Recent range shifts of moths, butterflies, and birds are driven by the breadth of their climatic niche. *Evol Lett*. 2023 Mar 12;8(1):89-100. doi: 10.1093/evlett/grad004.

da Silva CRB, Diamond SE. Local climate change velocities and evolutionary history explain multidirectional range shifts in a North American butterfly assemblage. *J Anim Ecol*. 2024 Aug;93(8):1160-1171. doi: 10.1111/1365-2656.14132.

Antão, L. H., Weigel, B., Strona, G., Hällfors, M., Kaarlejärvi, E., Dallas, T., ... & Laine, A. L. (2022). Climate change reshuffles northern species within their niches. *Nature Climate Change*, 12(6), 587-592.

Response:

Thank you for highlighting these relevant papers; we have added them to the revised manuscript [references #35, #36, and #37].

What time frame is this data from? The time span can have many impacts.

Response: The published studies and the expert assessments cover the period from 1991 onward (see line 706 and the supplementary methods).

Given the objective of including additional types of data, I found it surprising that the authors didn't look to see whether there was a difference in range shifts between English and non-English documents, or between expert assessments and from studies. There are more lessons to be learned from the new dataset.

Response: We thank the Reviewer for this excellent suggestion. We have now tested the importance of data integration and whether patterns differed between English and non-English literature. We have added the following paragraph in lines 326-340.

"To evaluate the importance of data integration from non-English language studies and expert assessments, we compared the composition of documented range-shift types across evidence sources using reference-weighted counts (i.e., unique source per range-shift type) (Extended Data 4). The composition of range-shift types differed among the categories of sources (Pearson's $\chi^2 = 38.14$, $df = 4$, $p = << 0.001$; Cramér's $V = 0.161$). Standardised residuals indicate that English-language sources over-represented horizontal range expansions and under-represented horizontal range contractions (expansion: 2.94; contraction: -4.14), and vice-versa for expert assessments (contraction: 5.73; expansion: -4.68), while there was no difference for non-English language studies (standardised residuals near zero). Together, these patterns indicate that reliance on English-language

studies alone would bias the assembled evidence base towards a pattern dominated by horizontal range expansions and that incorporating expert and multilingual sources reveals the broader pattern of dynamics across the globe, thus strengthening inference about the diversity of redistribution responses.”

Line 224- why are climate change impacts considered ‘indirect’? There are many ways of thinking about ‘indirect’, please define

Response: We have deleted both ‘directly’ and ‘indirectly’ from the sentence for clarity [Line 242].

Line 230-231- Elaboration here needed

Response: We have added the following sentence on lines 250-253.

“When dispersal capacities are insufficient relative to the pace of environmental change, species may persist in increasingly suboptimal environments, generating a lag between species’ realised distributions and their shifting climatic suitability (i.e., climatic debt).”

Line 240-241- Butterflies are not under-represented taxa

Response: We have deleted ‘under-represented’ [Line 263].

Line 326-329- Not clear what evidence there is that the European reflects “both genuine changes and stronger detectability”... “whereas lower records... reflect data scarcity rather than absence of biotic response.” Why not compare the type of data between Tropics and Europe?

Response: We thank the Reviewer for their suggestion and very constructive comments. We now performed a region-stratified comparison of the documented number of species shifting their ranges across Europe, tropical countries, and non-European non-tropical countries, further partitioned by range shift type (horizontal expansion, horizontal contraction, elevational shifts). We have added the following paragraph in lines 403-414 and added a supplementary figure (Extended Data 5).

“To quantify whether the reports of range-shifts in our database differed across regions, we extracted the country centroid, grouped countries into Europe, the tropics, and non-European, non-tropical countries. Shift-type composition differed significantly among regions (Pearson’s $\chi^2 = 389.13$, $df = 4$, $p = < 0.001$; Cramér’s $V = 0.277$), driven by an over-

representation of expansions in the tropics and higher relative representation of contractions and elevational shifts in Europe. In Europe, documented records were relatively balanced across expansions (41.3%), contractions (37.4%), and elevational shifts (21.3%). By contrast, expansions dominated in the tropics (80.3%) and in non-European non-tropical countries (64.8%) (Extended Data 5). Importantly, elevational range shifts were rarely documented in tropical countries, highlighting a major gap in the current evidence base and reinforcing the need for integrated monitoring approaches that improve detection of under-reported range-shift types.”

Results

Line 265-267- Was the diversity of the region taken into account with these estimates?

Response: Here, we refer to the overall results; therefore, we did not account for the ‘diversity of the region’. However, in the following section, when we compared the spatial patterns of range shifts, we did account for the regional species pool.

The results were not put into context with regional differences in temperature change.

Response: We thank the Reviewer for raising a very important point. To address the reviewer’s concern, we have added the following sentence on lines 515-518.

“Future work could combine standardised climate metrics (e.g., regional warming rates or climate-change velocity) and land-cover metrics to estimate the magnitude of range shifts and test whether the direction and rate of change align with regional exposure to environmental change.”

What are implications of differences in taxonomic coverage?

Response: We have added the following sentence on lines 347-350 to address this issue:

“Uneven coverage can bias global inference about the drivers and consequences if well-studied families disproportionately shape the evidence base; conversely, low-coverage families represent key monitoring gaps in which range shifts may be underdetected.”

Spatial patterns- What is the ecological relevance of the defining spatial patterns based on geopolitical boundaries?

Response: We agree that geopolitical boundaries are not ecological units and that species redistribution is structured by environmental gradients rather than national borders. We

used countries as the primary reporting unit for three reasons. First, the majority of sources in our synthesis report range shifts at the country level (often via national atlases, checklists, or because of national expert assessments), making the country level the most consistently extractable spatial unit across evidence types. Second, the national boundaries align with widely used denominators (e.g., national species pools) and are directly relevant to conservation planning (e.g., the Kunming-Montreal Global Biodiversity Framework targets), reporting, and resourcing, which are central to our aim of identifying global monitoring gaps.

To better explain our initial motivations behind working at the country level, we have added the following sentence on lines 711-713:

“We used countries as observational units because most sources report species range shifts at the national level (atlases/checklists/expert assessments) and because national baselines are required for conservation reporting and resource allocation.”

Further, results are mixed and could be better explained. First describe, what % of country’s regional pool have been assessed, then describe percentage of expansion, contraction.

Response: We have added the following sentence on lines 389-393 to start this section with the regional scenario.

“We then examined the composition of documented range-shift types to avoid conflating differences in the volume of evidence with differences in range shifts. Across continents, horizontal range expansions were the most frequently documented range-shift type, whereas the proportional contributions of horizontal range contractions and elevational range shifts varied among regions.”

Line 352-385- without “formal analyses” as the authors themselves acknowledge, I don’t find the section on drivers of range shifts very compelling. How did authors evaluate the relative importance of threats?

I appreciate the aim of incorporating more types of data but more rigour is needed when presenting them as integrated results. The authors lean on single species examples.

Response: We agree that our original wording could be read as evaluating the relative impact of threats, which would require comparable quantitative effect sizes or consistent multi-driver modelling across studies. Our synthesis does not aim to estimate the magnitude of causal effects. Instead, we extracted the driver(s) reported by each source for each documented range shift and mapped them to standardised IUCN threat categories. We therefore quantify the relative prevalence in the published evidence base (i.e., how frequently each threat category is reported as linked to documented range shifts), rather than ‘relative importance’ in the sense of effect size or causal dominance.

To strengthen rigour and address differences in reporting breadth across studies, we now present both raw prevalence (counts of unique reference–species–threat links) and an effort-adjusted prevalence metric that normalises each reference’s contribution by the number of threats it reports (each reference contributes a total weight of 1 across all threats, so studies listing many threats do not dominate). We also provide reference-level summaries and sensitivity analyses restricted to references reporting ≥ 2 threats. These changes are reported in the manuscript and shown in Extended Data 7. We have also revised the main text to replace causal/impact language and to clarify that reporting frequency may reflect both research attention and ease of quantification, not necessarily ecological effect size [e.g., 465-467, 469-471].

We have added the following paragraph on lines 455-465:

“These patterns may reflect both ecological differences and differential measurability and study emphasis among drivers (e.g., climate covariates are often more readily available than local pollution or fine-scale habitat degradation). However, an effort-adjusted sensitivity analysis produced qualitatively similar driver patterns to those based on raw counts, for both species-weighted and reference-weighted summaries (Extended Data 7). Where rankings differed between summaries, the differences reflected whether a threat was concentrated in a small number of species-rich references or distributed broadly across references, rather than evidence of comparative effect magnitude. Overall, climate change and severe weather is the most frequently reported threat category linked to butterfly range shifts in our database, and this ranking is robust to normalising for differences in the number of threat categories reported per reference.”

The single-species examples are simply to let the readers know the status of butterflies and how individual threats are impacting them.

Discussion

I found the discussion lacked substance and depth. Instead, it was a high-level summary of some mitigation strategies.

Response: We are grateful for this feedback. In this revised version, we have substantially revised this section. For example, we have added the following details.

“The fact that many species in our dataset exhibit only one range-shift type, whereas a minority shows multiple range-shift types across regions, further emphasises that species redistribution can be multidirectional and context-dependent rather than uniformly poleward.” [Lines 524-528]

“In practice, this means that even where climate becomes suitable, colonisation may fail if host plants and suitable microclimates are absent or if habitat connectivity is insufficient to bridge distant and isolated habitat patches.” [Lines 534-536]

“Overall, our synthesis indicates that conserving range shift butterflies requires two interconnected strategies: (i) reducing exposure to major drivers (e.g., climate and land-use change) and (ii) maintaining the ecological and landscape conditions that enable

populations to persist and move (e.g., host plants, microrefugia, and habitat connectivity). Minimising these threats requires an integrated approach that links climate mitigation, habitat protection, sustainable land management, and coherent policy action to prevent biodiversity loss in a rapidly changing world.” [Lines 575-581]

I didn’t find the framework to be particularly novel or helpful, and it is disjunct from rest of paper. For example, no ideas were provided for how to actually integrate citizen science data and local knowledge.

Response: We have now substantially revised the framework section to (i) link each step directly to the key gaps and biases identified in our results, and (ii) add concrete, implementable guidance on how citizen science and local knowledge can be validated, standardised, and integrated to support robust inference and conservation action.

Specifically, we have added the following information.

“Our semi-quantitative synthesis therefore highlights a dual challenge: (i) closing spatial and taxonomic gaps in baseline monitoring, and (ii) improving how heterogeneous evidence (i.e., including citizen science and local knowledge) is validated, standardised, and integrated so it can support robust inference and conservation action.” [Lines 587-591]

“However, these data are often opportunistic and heterogeneous in effort, making them difficult to use alone for robust inference about temporal trends, the rate of range change, or the mechanisms underpinning species range shifts. For this, citizen science is most powerful when paired with structured monitoring (Step I): standardised protocols provide consistent effort through time and space, improving range-shift detectability, and generating comparable baselines against which opportunistic observations can be calibrated¹⁴.” [Lines 631-637]

“We show that both horizontal range contractions and elevational range shifts are under-documented in many regions, yet these range-shift types are the most immediately relevant for extinction risks and protected-area adequacy.” [Lines 654-656]

Revisions for “Extensive climate-induced range shifts in butterflies across the globe”

We thank the Editor and the two Reviewers for their helpful comments. We have now revised the manuscript accordingly and addressed all remaining comments from Reviewer #1. Below, we detail our responses (in blue) to the Editor and Reviewer #1 while explaining the changes we made to the manuscript

(MainManuscript_revised_trackChanged; in red).

REVIEWER 1

I thank the authors for addressing most of my concerns in this revised version of the manuscript. I still have a number of smaller concerns that should be able to be addressed in another revision.

Response: We thank Reviewer #1 for their positive assessments and for providing further suggestions for improvement. We have carefully addressed each remaining comment and have revised the manuscript accordingly.

Note that in the merged file sent to reviewers, there is no version of the manuscript with the changes highlighted or tracked, so it's difficult to assess where in the manuscript changes were made.

Response: We are sorry to hear this. We submitted both the track-changed version and the no-changes-marked version. We have no idea why our track-changed was not available to Reviewer #1 and we sincerely apologise for this issue. For this new round of revisions, we also provide two versions of our manuscript, including the version with track changes, and hope the Reviewer will be able to access it this time.

I think the authors have made it clear throughout most of the manuscript that any rates reported are not the true global prevalence of range shifts (and drivers) across all butterflies, and have instead framed them as the relative prevalence of documented shifts within their gathered evidence base. However, this has not carried into the abstract where I think it could be mentioned there. For much of it, it still reads as an evaluation of the true prevalence vs. the relative reporting rates.

Response: We thank Reviewer #1 for catching this and we are sorry for missing it during the former round of review. We have now revised the abstract as well to clarify that our results reflect the relative prevalence of documented range shift types and reported drivers within the evidence compiled here, rather than absolute global rates across all butterfly species [e.g., Line 206, 211].

Abstract: “Most species’ (79%) range expansions were associated with climate change and severe weather” <- the placement of the 79% in the sentence: is that saying 79% of species

expanded their ranges or of the range expansions documented, 79% were associated with climate change and severe weather?

Response: To avoid any confusion, we have revised the sentence; it now reads as:

“Most species (80%) experienced range expansions, and most range shifts (79%) were associated with climate change and extreme weather events.” [Lines 210-211]

Abstract: “We find stronger support for range expansions and contractions across tropical countries, with less evidence for elevational shifts” <- stronger compared to what? Compared to non-tropical or temperate regions?

Response: For clarity, we have changed ‘stronger support’ to ‘many reports’. [Line 214]

L413-414: is this only horizontal contractions/expansions or poleward as well?

Response: This includes all types of horizontal range shifts. We define range shifts as partial changes in species distributions along the latitudinal and longitudinal dimensions (i.e., horizontal range shifts) (Table 1).

L347-350: Is there any reason to believe though that well-studied groups will show different patterns in shifts overall than lesser-studied taxa? Seems to me like it would only be a problem if one is not representative of the other, and the authors don’t present any reason a priori that one taxonomic group might have different sensitivity/exposures compared to others.

Response: We thank Reviewer #1 for raising this point. We have clarified that our concern is not that well-studied groups are necessarily biologically unrepresentative, but that uneven study effort among butterfly families may make species range shifts more likely to be detected and reported in some families than in others. [Lines 313-317]

L354-356: 80% + 19% <- what is the other 1%?

Response: The decimal was slightly above 80% (i.e., 80.95), but we rounded it at 80%. We have now corrected this wrong rounding and changed it to 81%. [Line 360]

L363: Why did the authors need to extract the country’s centroid if they were just grouping countries into Europe, the tropics, and non-European, non-tropical countries?

Response: Country centroids were extracted only to provide a simple, reproducible basis for assigning countries to the tropical versus non-tropical category in this broad regional comparison. We have added the following information in lines 370-375.

“We used a centroid-based approach to allocate each country to four groups (Europe, the tropics, non-European countries and non-tropical countries). Because several large

countries (e.g., the USA, Australia) span both tropical and temperate zones, as an example here, the USA was allocated to the non-European countries group.”

L389-392: “The density of European records reflects both genuine changes and stronger detectability of range shifts given long term biodiversity monitoring programs, whereas the lower density of range-shift records from the tropics likely reflects data scarcity rather than an absence of biotic responses.” <- I find this a bit of a reach; the authors cannot separate the detection vs. true prevalence of range shifts in your data, so what is the evidence behind this statement?

Response: In view of this comment and to avoid any confusion, we have decided to delete this sentence. [Lines 402-405]

L397: “An important note is that many tropical studies, range-shifts were simplified as expansion or contraction evidence when species newly appeared or disappeared locally (i.e., by comparing checklists rather than relying on quantitative data and proper statistical tests; Table 1), obscuring patterns in these regions. ” <- I think there is a word missing in this sentence, between “that” and “many”; perhaps “in”? Secondly, I could see how that would be an overall problem, but it doesn’t obscure your own patterns, correct? Because you did not extract any of the quantitative data from these studies (just the type of range shift). So this statement seems oddly phrased or placed, because the previous sentence is about your specific study.

Response: We thank Reviewer #1 for spotting that mistake. We have revised the sentence for clarity. [Lines 410-417]

“An important note is that in many tropical studies, range shifts were recorded simply as local fluctuations of population sizes or local population appearances and disappearances, rather than quantified using standardised analyses of range shift rates (Table 1), highlighting that despite our efforts adequate detection and reporting is scarce outside of Europe.”

L432: “The relative importance of different threat categories varied across range-shift types” <- I am not sure that the words “relative importance” are appropriate because (as the authors themselves state) the study cannot investigate the magnitude of impact or causal dominance of each of the threats. Suggest rewording to something like “the relative reporting rates”.

Response: Following Reviewer #1’s suggestion, we have changed it to ‘relative reporting rates’. [Line 449]

I wonder if for Figure 4B, this should be a relative proportion rather than absolute numbers. I worry that absolute numbers obscure: a) the difference in the sizes of species’ pool (i.e. more species in the tropics); and b) the number of assessments per continent. Low numbers of documented threats like in Africa and South America may just reflect a lower number of studies that examined threats rather than actual prevalence of threats.

Response: Following Reviewer #1's suggestion, we have revised Fig. 4B to show the relative proportion of reported threat attributions within each continent instead of the absolute numbers.

L432-436: What was the most frequently reported threat for species with documented range expansions?

Response: We now clarify in lines 454-455 that climate change and severe weather was reportedly associated with horizontal range expansions of 163 species.

L662-664: "an information that was either reported by a given study or that was assessed directly by our set of butterfly experts who relied on field observations" <- I think there is a typo in this sentence around "an information"

Response: We have deleted 'an' to correct this sentence. [Line 694]

L659-667: "Once each expert agreed, we asked them to share the same type of range shift information and to fill the same form by providing the type of range shift (i.e., distinguishing between horizontal range expansion, horizontal range contraction and elevational range shifts) (see Table 1), an information that was either reported by a given study or that was assessed directly by our set of butterfly experts who relied on field observations. Specifically, the experts compared the area of occupancy maps covered by field observations (see Table 1) to determine whether there has been a change in the species range and which type of range changes among the three studied range-shift types." <- I am confused about the expert elicitation process throughout the manuscript. When studies were provided to an expert, how were they matched to an expert? Why did they need an expert to judge rather than one of the regular data extractors. When it was "assessed directly by our set of butterfly experts who relied on field observations", was observation data provided TO the experts or BY the experts? And they based their judgment solely off of that, and no prior knowledge of the species? How were studies or species selected for each expert? Finally, the second mention of "see Table 1" seems misplaced as it has nothing to do with that section of the sentence.

Response: We thank Reviewer #1 for raising this important comment. Experts were not asked to evaluate any studies previously extracted by "independent" data extractors. Rather, based on their regional and taxonomic expertise, they were invited to provide independent assessments of butterfly species range shifts. These assessments were based on each expert's long-term field experience and knowledge [and, if true: including unpublished records, local atlases/checklists, or monitoring data], which they used to classify whether a species had undergone horizontal expansion, horizontal contraction, or an elevational shift. The reporting of these records was done using the same standardised template as for the literature review. We have revised the methods section of our manuscript to clarify how butterfly experts were selected, what information they contributed, and how their assessments differed from the literature-extraction workflow.

We also removed the misplaced Table 1 citation. We have revised this section in lines 696-705.

L676: Reference #30 is not a taxonomy; is this the right citation here?

Response: We apologise for this reference numbering error and we thank Reviewer #1 for spotting it. We have corrected it in line 712.

L676-677: From this wording, it seems like range shift information was extracted from the original study and this reference was used to extract country and country-level distribution data. I don't understand why country wasn't extracted from the original study/observation itself, and why it needed to be done in a posthoc manner like this. What if a species spanned multiple countries? Was a range shift counted in each even if the original study only documented it in one country?

Response: We agree with Reviewer #1 that the original wording was unclear. Country information was extracted directly from each study or expert assessment, as stated earlier in the Methods section [Line 676]. The final sentence referred only to taxonomic harmonisation: we used a recent global butterfly taxonomy to standardise species names, then matched those names to the external country-level distribution dataset used to derive national species pools for Figs. 2 and 3. We have now clarified this issue in detail in lines 712-716.

There is no Supplementary Figure S1 (cited in the Supplementary Material)in any of the documents provided to reviewers.

Response: We thank Reviewer #1 for spotting it and we apologise for this error; we have now revised it in the supplementary method section.

REVIEWER 2

I appreciate the extent of revisions made by the authors. I am content with the revisions made and am happy to recommend the manuscript be published.

Response: We thank the Reviewer for their very positive assessments and their encouraging comments. This is highly appreciated.

Revisions for “Extensive climate-induced range shifts in butterflies across the globe”

We thank the Editor and the reviewer for their helpful comments. We have now revised the manuscript accordingly and addressed all remaining comments from Reviewer #1. Below, we detail our responses (in blue) to the Editor and Reviewer #1 while explaining the changes we made to the manuscript

(MainManuscript_revised_trackChanged; in red).

REVIEWER 1

I thank the authors for their continued diligence in addressing my concerns. Just a few additional minor comments at this point.

Response: We thank Reviewer #1 for their positive assessments and for providing further suggestions for improvement. We have carefully addressed each remaining comment and have revised the manuscript accordingly.

L210-214: “Most species (80%) experienced range expansions, and most range shifts (79%) were associated with climate change and extreme weather events. A substantial proportion of species in our dataset contracted their ranges (27%) or shifted along elevational gradients (22%).” <- are all of these numbers reported in the main text? I could not find the 27%, 22% or 80% figures in the main text.

Response: We thank the reviewer for pointing this out. This information is discussed in lines 278-280.

“We gathered records on horizontal range expansions for 1427 species from 97 countries (Fig. 1B), horizontal range contractions for 480 species from 27 countries (Fig. 1C), and elevational range shifts for 381 species from 18 countries (Fig. 1D).”

L276: “Our dataset comprises 6182 range shift reports” <- does this number include instances of negative findings (i.e. species who did not experience a range shift)? If so, would suggest using a different term here than “range shift reports” as it may imply only positive instances. Similar feedback for L280-282. Do these numbers include instances of absences of horizontal range expansions, etc. or are they only the number of records with positive evidence?

Response: We thank Reviewer #1 for raising this important point. The numbers reported in this section refer only to positive instances of documented range shifts and do not include null findings or explicit reports of no change. Null results could not be compiled consistently across the heterogeneous evidence base, because many studies and expert-based sources reported only observed shifts rather than systematically recording species for which no shift was detected. We have therefore, in the last revision, revised the terminology throughout the manuscript to refer to documented records of range shifts.

L312-313: “Relative to the total number of known species per family, nymphalids had the highest percentage of species shifting their ranges, while riodinids had the lowest” <- wouldn't it also make sense to make it relative to the number of reports in your dataset, which would alleviate concerns about uneven effort between taxonomic groups?

Response: We thank the reviewer for this helpful suggestion. Our intention here was to assess taxonomic representativeness of the compiled evidence relative to the known global richness of each butterfly family, which is why we expressed values relative to the total number of described species per family. Using the number of reports in the assembled dataset would not alleviate uneven effort, because report counts are themselves strongly influenced by study intensity and repeated reporting within well-studied families. To address both perspectives, Fig. 1E already presents family-level patterns relative to both total known family richness and to the species represented in our assembled dataset.

L313-317: I appreciate the authors' clarification here. In their response, they address how uneven effort can affect detection and reporting of range shifts between groups. I would note that having uneven sampling effort between families would only underestimate the prevalence of range shifts in the understudied family if the denominator was the total number of species in a family, rather than the total number of reports in a family (see above comment). However, it is the word “bias” that still stands out to me here in this section, as well as the link to the actual drivers of range shifts (not just prevalence): the authors claim that “uneven taxonomic coverage can bias global inference about the drivers and consequences of species range shifts”. However, this would only introduce a bias if one family is not representative of the other, and the authors don't present any reason a priori that one taxonomic group might have different sensitivity/exposures compared to others, no a priori reason that this would bias drivers.

Response: We thank the reviewer for this helpful comment. We agree that our original wording was too strong. Our intention was not to imply that well-studied families are necessarily biologically unrepresentative of lesser-studied families, nor that uneven family-level coverage necessarily biases inference about drivers. Rather, our point is that uneven taxonomic coverage limits the taxonomic completeness and generality of the current synthesis, because well-studied families contribute disproportionately to the available evidence, whereas low-coverage families remain important monitoring gaps in which comparable range shifts may be underdetected. We have revised the text accordingly.

“Uneven taxonomic coverage can limit the taxonomic completeness of this synthesis, because well-studied families contribute disproportionately to the available evidence, whereas low-coverage families represent important monitoring gaps in which comparable species range shifts may remain underdetected.” [Lines 314-318]

L373-375: “Because several large countries (e.g., USA, Australia) span both tropical and temperate zones: as an example here, USA was allocated to the non-European countries group.” <- the sentence structure here is awkward and grammatically incorrect.

Response: We thank the reviewer for pointing this out. We have revised the sentence.

“Because several large countries (e.g., the USA, Australia) span both tropical and temperate zones, we classified them following our regional groupings; for example, the USA was allocated to the non-European countries group.” [Lines 366-369]