Dear authors,

first I'd like to apologize for the time it took me to make this first decision -- waiting for a third review that never came is a poor excuse, so mea culpa.

After reading the comments of both referees, I'd be happy to recommend a revised version of your paper.

The idea and the content have been deemed very useful by both reviewers (and I completely agree with them). However, some interesting/needed bits are buried in the Appendices and it would be better if you could re-organize the contents so that it becomes easier to access these.

Beside minor corrections and some editing of language errors, I think you should focus your revision of the paper on (1) restructuring some parts (and recombining parts of the main text with summaries of the Appendix tables) so that information comes more easily to the curious but rapid reader and (2) rephrasing your conclusions in order to be very precise about the consequences of your findings for already published papers that use the InsectChange database (i.e. make a clear difference between cases where you have proof the published results are no longer valid vs. cases where the amount of errors very strongly suggest that this could be the case, but you haven't checked explicitly). It is fair to assume that a poorly curated database will probably lead to bad results, but it doesn't disprove the findings reported in earlier studies unless you actually redo the analysis and show it does not hold any more.

You will see that the reviewers' reports are very constructive and bring some useful tips as to how revise your work.

Thanks again for producing this important paper. I am looking forward to reading the revision of your manuscript.

Sincerely,

François Massol

*by **Francois Massol**, 25 Mar 2024 15:18*

Manuscript: **https://doi.org/10.1101/2023.06.17.545310** version: 2

Montpellier, 20 July 2024

Dear Dr Massol,

We carefully read the comments of both reviewers as well as your suggestions and took all of them into account, which required quite a lot of work. They were very helpful, and we hope that the manuscript is now clearer and more convincing.

As requested, we restructured the text of the manuscript to make it more easily informative to rapid readers. We moved the former Appendix 1, which explained the types of problems and their possible impact on insect trend analysis, to the main text of the manuscript, in the form of a more synthetic Table 1. This Table also provides information on the frequencies of problem types that were previously presented in Figure 1. Figure 1 now shows how problem types are distributed between freshwater and terrestrial realms. The results of logistic regressions testing for a difference between realm for each problem type are shown in a new Appendix 1.

To meet the reviewers' expectations, we have brought two major changes to make a more convincing case for the extent to which analyses drawn from InsectChange data can be flawed, both in terms of assessing insect trends and identifying their drivers. The first addition pertains to freshwater datasets that have considered the whole invertebrate assemblage instead of just insects and where the insect share was calculable. We now show that the trend obtained when considering insects alone is significantly lower than that obtained when considering the whole invertebrate assemblage; this has been tested by a linear mixed model and is illustrated by the new Figure 4 and Appendix 2. The second addition is a more precise account of the major methodological problem caused by the inadequate matching of InsectChange data with external databases at the local scale. We have now precisely assessed the inadequate assignment of geographical coordinates to insect sampling points in InsectChange for a matching with the local scale ESA-CCI database from which the agricultural and urban land cover were derived; this is illustrated by the new Figure 6, which explains the different causes of misassignment. We now classify the inadequate assignment of geographic coordinates as a methodological issue rather than an error, as was previously the case. We also now present boxplots of the minimum errors in overestimated or underestimated cropland cover assessments, which required an additional important work as this information was not formerly included in our Appendices.

Moreover, when we went back into the database to answer the reviewers' questions, we found new problems. First, we found more inadequate geographical coordinates, which are now 112 instead of 88. We also showed that the time series used in the 20 plots of the dataset corresponding to Study 1473 should be discarded, as the sampling method changed over time in the original study. We illustrated this problem with the new figure 2f, showing that considering or not the first year of record can radically change the temporal trend in the source dataset. Finally, we realised that abundance or biomass data were not standardised in InsectChange, i.e. not systematically expressed in international units. We argue that the $\log_{10}(x+1)$-transformation of these heterogeneous data, as has been done in the *Science* article by van Klink et al. (2020), biases the comparison of temporal slopes between studies and the estimation of an overall trend. We wrote a paragraph on this critically important methodological issue.

We agree that we did not formally disprove the findings reported in earlier studies given that we did not redo the analysis. As we argue, a main problem with InsectChange is that it gathers datasets from studies with differing aims, most often taking the abundance or biomass numbers from these studies without accounting for the internal drivers of insect change that were addressed in these studies, and that it is not representative of worldwide habitats in terms of insect groups

and anthropogenic changes. Redoing the analysis in a meaningful way would mean completely overhauling the database to make it more representative of the diversity of habitat conditions and associated insect abundances worldwide, not only correcting some numbers. We believe that correcting and consolidating the database and reworking analyses should be the task of the authors of the database themselves, not ours. Nevertheless, in order to make our criticism more constructive, we now present several ways of improving the selection of data to make them more representative of habitat conditions and insect numbers on a global scale. (lines 355-364 p12). As recommended, we have been careful not to be so assertive about the quality of the results already obtained from InsectChange when we did not prove that our recommended changes would have modified the results.

As regards to English, we had the manuscript proofread (we can provide you with a AJE certificate if you wish). We have also incorporated the other minor comments made by the reviewers.

We downloaded the eight supplementary files, including now the new compressed file *Stat_invertebrates.zip* gathering data, scripts and outputs of our analyses, in the Figshare repository, where the data will be permanently archived if the manuscript is accepted for publication (doi: 10.6084/m9.figshare.23458877). If the manuscript is accepted for publication, a hyperlink to the repository doi will be added to each mention of these supplementary files in the text. The private-for-peer review link for these supplementary files, is the following: https://figshare.com/s/c7e8359aa65b6830580a.

We hope that the changes made to the manuscript will meet your expectations.

Sincerely yours,

Laurence Gaume and Marion Desquilbet

Review by anonymous reviewer 1, 12 Feb 2024 15:29

Dear reviewer,

We would like to thank you for your helpful comments on our manuscript submitted to *PCI Ecology*. We did our best to incorporate your suggestions and to address the concerns raised in your report. In what follows, we outline how we addressed each of your points. For clarity, we have copied each of your comments and indicated our answer below.

Summary

This paper by Gaume and Desquilbet is an important piece of work and I recommend publication after minor revision.

Gaume and Desquilbet take a deep dive into the InsectChange database that has been assembled by other researchers and used to publish several peer-reviewed papers in high-profile journals like *Science*. Some of these papers have claimed to find trends in insect abundance that run counter to existing knowledge. For example, authors who have used the InsectChange database have published meta-analyses claiming that while terrestrial insect abundances are in decline, freshwater insect abundances are increasing. The latter trend runs counter to a considerable body of prior research.

Inferences from meta-analyses are only as good as the data they are based on (e.g., garbage in-garbage out). And datasets like InsectChange are only useful when the information contained in them is accurate, reliable, and appropriate for testing the proposed hypotheses.

In the current paper, Gaume and Desquilbet show that the InsectChange database is riddled with data errors, inconsistencies, methodological problems, and information gaps showing this dataset did not receive the quality assurance/quality control needed for use by scientific studies. The authors go on to show how these errors, inconsistencies, problems, and information gaps undermine results of publications that have used this database and cast doubt on their conclusions. As just one example, conclusions about freshwater "insects" from analyses of the InsectChange database appear to be incorrect, as they are driven by increasing abundances of non-insect invertebrates that tend to increase as water quality declines (e.g., oligochaeta, turbellaria, amphipoda). These non-insect invertebrates were apparently added to the dataset by accident, as those assembling the InsectChange database did not properly differentiate between insects and other types of non-insect invertebrates.

I am highly supportive of publication of the critique by Gaume and Desquilbet. It took a huge amount of work to review the quality of InsectChange, and the authors have done the scientific community a huge service in doing so. I wish more databases like InsectChange received the scrutiny they should prior to being used for numerous publications.

I hope my suggestions below are helpful:

R1: Editorial work. The manuscript contains several English language errors and would benefit from 1more editing before publication. I kept track of some issues during my reading of the first part of the paper, and mention these below under 'Suggested edits'. However, there were enough editorial errors that I stopped noting them about a third of the way through the paper so that I could focus on the more substantive issues. I do want to emphasize that none of the editorial issues affect the analyses, conclusions, or message of this paper. But the paper would be a bit easier to read with a tad more editorial work.

LG&MD: we had the manuscript proofread by both an English-speaking person and AJE editing services. We hope the language is better now.

R1: Reliance on examples. In Figures 2-4, the authors rely on use of select examples pulled from the InsectChange dataset to illustrate their points. Yet, it is not always clear whether these select examples are broadly representative of the dataset as a whole. For example, the text starting line 72 reads "Among the errors, 35 datasets considered taxa other than insects, arachnids or entognaths (hereafter collectively referred to as "insects" for brevity), most often including the entire invertebrate assemblage instead of insects only, sometimes biasing the trends to the point of reversal (Figure 2a)." Figure 2a might lead the reader to believe that these errors in 35 datasets alter conclusions about time-series. However, upon closer inspection it looks as if the authors have chosen just one plot from a singles study to illustrate the problem (Study 1435 – Plot 448).

There are other places in the text / figures where it is clear that Gaume and Desquilbet have assessed the entire InsectChange database and concluded that errors are pervasive (e.g., Figure 3b). Even so, the text could be more clear about when select examples are being used to illustrate a problem vs. when analyses are illustrating a collection of errors (i.e. 35 datasets where non-insects were considered) that lead to conclusions different from the original papers.

LG&MD: We have now clarified in the text that former Figures 2 and 4 (now Figures 2 and 5) show examples of problems found in InsectChange, and are intended to show the diversity of problems encountered in the database. We also specified, for example in the caption of Figure 2, that they affect the trends of insects (sometimes actually invertebrates) **in the original time series** and not the overall trend, as now also stated in the sentence line 110 p 5 "*35 datasets … most often included the entire invertebrate assemblage instead of insects only, sometimes changing the insect trend of the original time series to the point of reversal (e.g., Figure 2b, details in section 2)*". It is the accumulation of these types of errors that can have an effect on the estimation of an overall trend, as mentioned in the end of the introduction, lines 68-70 p. 2: "*Our analysis highlights numerous limitations in the constitution of this database, the accumulation of which is likely to bias any assessment of insect change and drivers of change*".

In support of this last sentence, we now provide evidence for freshwater datasets that considering the entire invertebrate assemblage rather than just insects has an impact on overall trend estimations. More specifically, we show that in the data subset gathering datasets where it was possible to calculate the insect share in assemblages, a significantly lower and negative biomass trend is obtained when considering only insects than when considering the assemblage as a whole (see the new Figure 4 and lines 259 to 301 pp. 10-11) and we show that this result holds when biomass and abundance data are mixed.

As for the former Figure 4 (now Figure 5), we specify in the legend that graphs a to f are examples illustrating studies analysing the impact of specific factors and graph g shows how studies of these types are distributed within the freshwater database and the terrestrial database. We hope that this clarifies that we are showing examples and their representativeness in the database.

R1: Problems and their impact. It was not always clear to me from the text how select problem types shown in Figure 1a might impact conclusions from the InsectChange database. For example, I did not initially understand how problem type 1 – Inadequate geographical coordinates -- might influence conclusions from the database. It wasn't until I read Appendix 1 describing the consequences and risks of each problem type that it became apparent to me.

The manuscript seems to rely heavily on readers going into the Appendices to get more information before they can understand the main body of the text. The authors may want to consider taking essential elements out of the Appendices and putting them into the main body of the paper to make the reading easier. For example, Appendix 1 could be a Table in the main body of the paper … or, if that is too cumbersome … perhaps be more clear when describing the consequences of each problem type in the main body of text.

LG & MD: For clarity, we moved the earlier Appendix 1 describing the problem types and their possible impact on insect trend analysis to the main text of the manuscript, as a more synthetic Table 1. This Table also provides problem type frequencies, previously shown in Figure 1. We changed Figure 1 accordingly, so that it was not redundant but complementary with the information provided in Table 1. Figure 1 now shows how the problem types are distributed between freshwater and terrestrial realms and how they differ in frequency between the two realms.

We have also been careful to be more explicit in the text. We based our explanations on examples selected from Figure 2 and now hope that the manuscript explains more clearly each problem type and its consequences.

As for inadequate geographic coordinates, we now devote the entire sub-section section 4.1, lines 390-457, pp. 14-16 to this problem and its consequences for the analysis of the impact of agriculture, urbanisation or climate on insect trends when using external databases providing information at a local scale. After a more thorough examination of the geographic coordinates assigned in InsectChange to insect sampling points in the source studies, we now show that a total of 233 terrestrial geographic coordinates were inaccurate for a matching with the ESA-CCI database from which local agricultural and urban land covers were derived. This is illustrated in the new Figure 6. We also show more clearly that the local cropland cover is overestimated because of this inaccurate matching and because of misclassifications from the automated interpretation of satellite images in ESA-CCI, notably between croplands and grasslands, heathlands, steppes, barrens, prairies, shrublands, marshlands, natural vegetation areas, parks or golfs. This proves that InsectChange does not allow to study the local impact of agriculture on insect temporal trends and suggests that the same holds for the local impact of urbanisation and climate change. We therefore now classify the inadequate assignment of geographic coordinates as a methodological issue rather than an error, as was previously the case.

R1 : Suggested edits.

Line 22. Change "… allow one to study"

Line 23. Change to "… and extreme vigilance in use of the InsectChange database"

Line 34. Change to "… with trend assessments hampered by lack of data …"

Line 41. Change to "… than reported by prior authors, and they further proposed that the diversity of freshwater insects is increasing rather than decreasing."

Line 61. Change 'totalizing' to totaling.

LG & MD: The suggested edits in lines 22 to 61 have been implemented (in these edits, we rephrased "… than reported by prior authors, they further proposed that the diversity of freshwater insects is increasing rather than decreasing" to "than reported by previous authors, and further proposed that freshwater insects were increasing rather than decreasing", as van Klink et al. refer to abundance and/or biomass of insects, not diversity).

R1: Lines 74-77 (Figure 2b). I was unable to tell whether the sentence was referencing the blue line in Figure 2b (insects) or the red line (moths+beetles reported as insects). It was difficult to understand the error tilted 'Misreporting of select insect groups' based on the text and figure.

LG & MD: In the former Figure 2b (now Figure 2a), we now specify that the blue line refers to "light-trapped insects in source study" and that the red line refers to the selection by InsectChange of moths and beetles from the light-trapped insects in the source study. The insects selected were reported as "light-trapped insects", while they only corresponded to part of this group of insects. This is why we have titled Figure 2b "Insect group inadequately reported".

R1: Lines 79-80 (Figure 2f). Based on the text, I was not able to understand what was meant by 'datasets or plots had overlapping data'.

LG & MD: We have clarified what we meant by "overlapping data" in the text by taking an example illustrated in figure 2g (lines 135-145 p. 7): "*For example, InsectChange Study 1452, which is illustrated in Figure 2g, examined the change in biomass of the invertebrate assemblage*

after the creation of the Kama Reservoir in Russia. InsectChange Plots 456, 457 and 458 corresponding to the upper, central and dam sections of the reservoir, respectively, include data from 2003 to 2015 mainly for insects, and Plot 455, corresponding to the average sampling in the three sections of the reservoir, includes data from 1955 to 2013 on the entire zoobenthic assemblage. From 2003 to 2013, insect data from Plot 455 therefore overlap with invertebrate data from Plots 456, 457 and 458."

R1: Lines 80-82. It was unclear to me why erroneous or imprecise geographic coordinates in 88 datasets might lead to inaccurate conclusions about insect time trends.

LG&MD: Lines 80-82 of the former version of the manuscript were not meant to imply that the problem of inadequate geographic coordinates led to inaccurate conclusions about insect trends. We argued later in the text that it led to inaccurate conclusions about the drivers of insect trends. To be more explicit on this point, we have completely restructured the text and devoted an entire subsection (section 4.1, pp. 14-16) to the issue of inadequate geographic coordinates, which is now classified as a methodological issue. We have extended our analysis of the inadequacy of geographic coordinates and we now provide more detailed explanations and show how this issue affects the analysis of agriculture and is also likely to affect the analysis of urbanisation and climate as potential drivers of insect decline. We explain in the text (lines 391-400 p. 15) that "*The local scale around each plot is defined in InsectChange as the area of 900 m × 900 m centred on the 300 m × 300 m ESA-CCI cell encompassing the geographic coordinates assigned to the plot and including the eight surrounding ESA-CCI cells. This area is used to estimate the cropland or urban cover at the local scale. The adequacy of these local-scale indicators hinges on the premise that, for each plot, the geographic coordinates assigned to the plot in InsectChange are precise enough to point to the insect sampling area, and that this sampling area is included in a 900 m x 900 m square (hereafter referred to as a "local-scale square", Figure 6a). However, this was not the case for almost a quarter of the terrestrial plots (233 out of the 985 plots). This methodological issue affected 63 of the 103 terrestrial datasets included in InsectChange*". Therefore, the geographic coordinates assigned to InsectChange plots are often inadequate for indicating sampling locations and therefore analysing drivers of insect change at the local scale based on ESA-CCI information. We also added a new figure to provide a better explanation (Figure 6).

R1: Line 90. Should this line be referring to Figure 2e or a different panel?

LG & MD: It was referring to Figure 2f and not 2e, we are sorry for this mismatch and corrected it.

Please note that we made several other changes to the manuscript to meet the expectations of the other reviewer. The main one deals with freshwater datasets that have considered the whole invertebrate assemblage instead of only insects and where the insect share was calculable. We now show that for this datasubset of InsectChange the trend obtained when considering insects alone is significantly lower than that obtained when considering the whole invertebrate assemblage; this has been tested by a linear mixed model and is illustrated by the new Figure 4 and Appendix 2. In addition to the other supplementary files, the scripts and outputs of the analyses are available in the new R. project file, *Stat_invertebrates.Rproj* found in the *Stat_invertebrates.zip* compressed file with the private-for-peer review link, https://figshare.com/s/c7e8359aa65b6830580a. The data will be publicly available if the manuscript is accepted for publication.
You can also note in the main text and in Figure 3 that the mean percentage of insects and associated coefficient of variation slightly changed (mean = 48.7% instead of 48.5%, CV = 36.4%

instead of 34.8%) because for Study 1452, Plots 455 and Plot 459, we found new time records from which we could extract the percentage of insects (*FreshwaterNonInsects.xlsx*, sheet 3).

We hope that we have successfully addressed your concerns. Thank you for your helpful comments, which, we believe, enabled us to improve the quality of our paper.

Review by anonymous reviewer 2, 09 Feb 2024 07:31

Dear reviewer,

We would like to thank you for your helpful comments on our manuscript submitted to *PCI Ecology*. We did our best to incorporate your suggestions and to address the concerns raised in your report. In what follows, we outline how we addressed each of your points. For clarity, we have copied each of your comments and indicated our answer below.

R2: The manuscript "InsectChange: Comment" presents an exhaustive review of InsectChange, a large database of long-term time-series on insects, arachnids and entognaths' assemblages. The authors made an extraordinarily thorough work in scanning InsectChange (dataset per dataset). They detected (and reported on) data errors and inconsistencies that were inherited from other databases (e.g., BioTIME and Global Population Dynamics Database) or introduced by the processing of the original datasets. Importantly, the authors made the effort to compare the datasets included in InsectChange (in their current status) with the corresponding original data sources. Beside tracing back the origin of data errors and inconsistencies, this allowed identifying other methodological issues affecting InsectChange. Examples are the lack of information on whether data were collected under some experimental treatment or the undocumented inclusion of invertebrates other than insects in assemblages thought to consist only of insects.

I personally agree with the authors that some of the issues they found in InsectChange could potentially bias estimation of temporal change of insects. I believe that their review is an important basis of improvement of InsectChange. Considering (and addressing) all errors and inconsistencies highlighted by the authors will surely make InsectChange a better database to investigate insects' trends across wide spatiotemporal scales!

My main concern is not about the review itself, which, once again, is an important piece of work. Several times in the comment, the authors claim that the issues they found (fully) undermine or invalid previous findings based on InsectChange (often referring to the global analyses carried out by van Klink et al. 2020). I agree that addressing all errors, inconsistencies and other issues in InsectChange, and re-running analyses with a 'better' version of the database, would likely change the 'numbers' estimated in previous works (e.g., measures of effect size or estimates of temporal trend in insects' abundance). Yet, the authors did not explicitly test that in their comment. At least, I did not find a comparison between findings from previous studies and those obtained using an updated version of InsectChange. Rather, the authors made a large series of examples of datasets affected by data-related errors or methodological issues, and of the potential bias they could introduce when estimating temporal trends in insect change. However, these dataset-specific examples do not allow quantifying/guessing the real impact of the different issues on previous research (e.g., reversing of trends in insect change previously described in studies implemented at large spatial scales). This is just to say that I would personally avoid stating that errors found in InsectChange fully undermine previous research (as stated in the conclusions section), as the authors never explicitly quantified how much previous findings based on InsectChange deviate

from those potentially obtained addressing all errors and issues. I see this review as an invaluable basis for improvement rather than as a means to debate on the validity of previous work (at least as long as no explicit comparison on that is provided).

LG & MD: We agree that we did not formally disprove the findings reported in earlier studies given that we did not redo the analysis. We have reformulated the conclusion and now state that the problems we discuss "*call into question*" (rather than "*fully undermine*") the results obtained so far from this database. As we argue, a main problem with InsectChange is that it gathers datasets from studies with differing aims, most often taking the abundance or biomass numbers from these studies without accounting for the internal drivers of insect change that were the subjects of these studies, and that it is not representative of worldwide habitats and insect dynamics. Redoing the analysis in a meaningful way would mean completely overhauling the database to make it more representative of the diversity of habitat conditions and associated insect abundances worldwide, not only correcting some numbers. We believe that correcting and consolidating the database and reworking analyses should be the task of the authors of the database themselves, not ours.

In section 2, on the basis of the estimation of a mixed linear model, we now provide evidence that considering the whole invertebrate assemblage rather than insects may significantly influence the temporal trend. We showed this for the freshwater InsectChange datasets that have considered the whole invertebrate assemblage instead of only insects and where the insect share was calculable. We find that the temporal trend of insect change is significantly lower when considering insects alone rather than the whole invertebrate assemblage, for biomass and for abundance and biomass combined. This is illustrated by the new Figure 4. This shows that the specific problem of the inclusion of non-insects can have a significant influence on the estimation of "insect" trends at a wide scale.

R2: A somehow similar concern is that authors made several statements about the danger posed by some methodological issues affecting InsectChange (in the main text as well as in the appendices), but it was not always clear to me what was the real impact of these issues on the database. As an example, section #4 of the comment discusses the accuracy of local cropland cover derived from remote sensing data (ESA-CCI), which is provided in InsectChange as a ready-to-use driver to analyze insect change. In short, based on a comparison of cropland cover as reported in InsectChange with information from the original data sources (and the visual interpretation of satellite images), the authors conclude that ESA-CCI data are not adequate to estimate local cropland cover, as these data consistently under/over estimate it. Although the authors present a detailed comparison in the CroplandCover.xlsx appendix, they do not report in the main text what is the average difference between the cropland cover reported in InsectChange and the correct cover reported either in the original source or visually estimated from satellite imagery. I would suggest to provide (whenever possible) simple measures of the difference between data provided in InsectChange vs. original or alternative sources. This would greatly help readers to understand what is the real impact of a given issue on the database.

LG & MD: The average difference between the cropland cover reported in InsectChange and the correct cover was actually not assessed in our previous supplementary file *CroplandCover.xlsx*. The information available in the source publication and the satellite images generally do not enable to obtain precise cropland covers. We have extended our former analysis to include, when possible, an assessment of a minimum or maximum cropland cover per plot on the basis of the measurement of clearly identifiable parts of local land covers. This now enables us to show a minimum error in InsectChange estimates for overestimated and underestimated cropland covers (lines 473-477 p. 17), "*On the basis of only clearly identifiable parts of the land cover, we found that for 129 geographic coordinates for which assessment was possible, the assessment errors*

were very wide-ranging: the minimum overestimation of the cropland cover varied between 3% and 100% (mean: 45%, median: 36%) and its minimum underestimation varied between 1% and 67% (mean: 15%, median: 12%, Figure 7a2)" and Figure 7a2, p. 18). Calculating all these areas and adding this information, and our new detailed analysis of the inadequacy of geographic coordinates (subsection 4.1 pp. 14-16), have represented a very important additional work. Because of the time required, it is not possible for us to extend such order-of-magnitude calculations to all the issues we raise in the manuscript, although we do of course understand that they are very useful for understanding the real impact of a given issue on the database.

R2: List of detailed comments:

Line 67 (Figure 1 caption): what does the '*' refers to in the figure?

LG & MD: The '*' formerly indicated in Figure 1 and now indicated in Table 1 for the problem type 'Inadequate cropland estimation' specifies that the percentage was calculated only for terrestrial datasets since we did not assess the cropland cover for freshwater datasets.

R2: Line 70 (Figure 1 caption): does 'studies' mean the same as 'datasets' here? The terms 'studies' and 'datasets' are used interchangeably in the text, but it is not always clear whether they always mean the same thing.

LG & MD: We have replaced "studies" with "datasets" in Figure 1 caption. In the manuscript, we now more consistently use "studies" to refer to the original publications and "datasets" to refer to datasets created in InsectChange from these publications.

R2: Line 74: would it be possible to quantify (and report) how many times temporal trends were reversed if exclusively considering insects? This could provide a clearer idea of the potential effect of this error on the temporal trends estimated using InsectChange.

LG & MD: We now show in Figure 4a p. 10 that 7 out of the 21 plots where it was possible or relevant to assess an insect trend, i.e. a third of the plots, showed a reversed trend when only insects were considered.

R2: Lines 80-82: what does 'erroneous coordinates' exactly mean here? Does it mean that, for example, a plot located somewhere in Spain was instead erroneously located in Greece? Or rather that longitude and latitude were swapped? Similarly, I would clarify what is meant by 'not precise enough coordinates'? Does this refer to the number of decimal places of the coordinates pair? I believe these details are important to allow readers understand what is the problem with these data. Finally, it is not clear whether errors with coordinates were introduced by InsectChange or somehow inherited from the original data sources.

LG & MD: We have carried out a significant amount of work to assess more precisely why many geographic coordinates were inadequate and the extent of this inadequacy, as reported in subsection 4.1, pp. 14-16, including the new Figure 6, and in our supplementary file *CroplandCover.xlsx*. We now detail two different possible reasons for inadequate geographic coordinates: a sampling area too large for a local scale analysis or an inaccurate location of the sampling area. Our Figure 6 reports two boxplots showing the extent of these problems in InsectChange.

Our main text provides an example of sampling area too large for local analysis (lines 436-443 p. 16): "*For example, the 370 km distance is related to Study 1470, where InsectChange extracted*

a mean hymenopteran time series from Belarus in a unique plot and assigned it a location in Belarus where no sampling actually occurred. The information from the source study gave the names of the areas where the insects were sampled, allowing calculation of the distances between sampling points, which ranged up to approximately 370 km. Therefore, the local-scale indicators calculated around the geographic coordinates assigned to this unique "plot" are not meaningful for informing on the local conditions around the actual sampling points".

It also provides examples of inadequate plot locations lines 448-457 p. 16): "*For example, from the columns PlotName, Location and DetailsPlot in the table PlotData.csv of InsectChange, Plots 1656 (Study 1266) and 1670 (Study 1006) represent the Cairngorms site of the UK Environmental Change Network, but were inadequately assigned the geographic coordinates of the 450 km-distant Yr Wyddfa/Snowdon site. Other sources of inadequacy are detailed in our supplementary table CroplandCover.xlxs. They include the use of different geographic coordinates than those provided in the source study, an error when transforming geographic coordinates to the decimal format, and the inexact attribution of geographic coordinates in cases when they were not provided in the original study, and the use of geographic coordinates that were approximate or erroneous in the original publication or database from which they were extracted.*".

R2: Figure 2: In the figure caption, I would mention that 'blue' is used to represent insect change as obtained through a 'correct' use of the data, while 'red' is used to represent insect change as obtained using data in their current status. Also, the background of the plots could be colored to better discriminate between examples of errors and inconsistencies (for instance using the same color palette used in Figure 1a).

LG & MD: as suggested, the figure caption in Figure 2 now mentions: "*Examples of errors (blue blackground) and inconsistencies (orange background)*" and "Problematic insect dynamics are represented by red dashed lines, while non-problematic or corrected insect dynamics are represented by solid blue lines". We kept this red/blue colour code for the other Figures in the manuscript (e.g. new Figure 4) and in the online supplementary Appendices (e.g. *FreshwaterNoninsects.xlsx*, third sheet).

R2: Line 90: note that Figure 2e is about 'non consideration of sampling effort' (and not about 'inconsistencies of taxa between plots of a same dataset').

LG & MD: This has been corrected in the new version (p. 6).

R2: Line 114: note that the Excel file is named 'FreshNoInsects' and not 'FreshNonInsects'.

LG & MD: The excel file is now named *FreshwaterNonInsects.xlsx*. This file provides a new sheet that details the data subset, which we used to run the analysis on invertebrate trends. In addition to the other supplementary files, the scripts and outputs of the analyses are available in the new R. project file *Stat_invertebrates.Rproj* found in the *Stat_invertebrates.zip* compressed file with the private-for-peer review link, https://figshare.com/s/c7e8359aa65b6830580a. The data will be publicly available if the manuscript is accepted for publication.

R2: Figure S1 (Appendix S2): what does 'computable albeit not considered' mean in the title of Figure S1? Also, I would avoid referring to other appendices (or to the main text) to find details about a figure or table. As an example, in Figure S1 (Appendix S1) authors refer to the caption of

Figure 3 (main text) to read about what the colors of the pie chart are associated with. I strongly suggest to avoid that as the comment (main text and appendices) is very dense and one can get easily lost while looking for information across appendices.

LG & MD: We have changed the title of Figure S1 as follows: "*Insect share (mean percentage) in time series included in InsectChange relating to whole invertebrate assemblages and where it was possible to calculate the insect share and its variation over time (standard deviation), (A) for all time records for 37 plots from 6 studies or (B) for some available time records for 11 plots from 7 studies*". We annotated the pie chart to facilitate its understanding. We hope this is now clearer.

R2: Line 122: it should be 'in almost half of the plots' here, right?

LG & MD: Right, we added "almost".

R2: Figure 3: not entirely clear to me what 'insect % sometimes computable' means? Also, I do not understand what 'insects inferred to be dominant' means.

LG & MD: as we now specify in the figure caption (p. 9), "Insect % sometimes computable" means that it was possible to extract the percentage of insects for some records of the time series, as was specified in Table S1 of Appendix S2. Note that the mean percentage of insects and the associated coefficient of variation slightly changed (mean = 48.7% instead of 48.5%, CV = 36.4% instead of 34.8%) because for Study 1452, Plots 455 and Plot 459, we found new time records from which we could extract the percentage of insects (*FreshwaterNonInsects.xlsx*, sheet 3). As for the data subsets where insects were 'inferred to be dominant", in these subsets, the percentage of chironomids, which are part of the insects in these InsectChange-selected data subsets of original datasets, could be calculated for each time record and was most frequently well over 50%. We now specify this in Table S1 of Appendix S2 and in the caption of Figure 3. The mean percentage of chironomids over time records / plots / metrics for the datasets 1448, 1449, 1451 and 1457 were actually 77%, 71.6%, 72.9% and 65.6%, respectively.

R2: Lines 135-146: I understand (and agree) that 'internal drivers' (i.e. those originally considered/imposed in the source studies) may better explain insect change than 'external drivers' (i.e. those derived from external databases and not necessarily linked to the experimental, study-specific context). However, I am wondering how (and if) one could really account for all study-specific contingencies in analyses focused on large spatiotemporal scales, such as that carried out by van Klink et al. (2020). Should we refrain from doing large scale analyses if we can't control for all local-scale drivers possibly affecting context-specific changes? Should we break the database in small pieces to analyze separately and then try to put together inferences drawn from these single analyses? Isn't the whole idea of large scale analyses about finding emerging patterns rather than explain local contingencies?

I am saying this as the authors (rightfully) mention several case-specific studies where trends of insect change deviated from those expected by the original authors of the study (or from those reported by van Klink et al.). However, I am wondering how these issues should be accounted for when analyzing all data together with the aim of finding general/global trends of insect change (and/or possible correlates of the change to make predictions).

LG & MD: we have included a longer discussion of these issues lines 323-349 p. 13, were we state: "*This analysis raises the question of whether the data included in InsectChange are representative of habitat conditions and associated insect abundances worldwide, particularly in*

freshwater. While the selection of data according to specific and consistent criteria is a necessary condition for a meta-analysis to lead to robust conclusions (Englund et al., 1999), it was not met in InsectChange. The inclusion of time series with specific experimental designs to address ecological questions with differing purposes and expectations raises three issues for meta-analyses and other syntheses carried out using this database. (1) Such inclusion does not fit the definition of a meta-analysis as "a set of statistical methods for combining the magnitudes of the outcomes (effect sizes) across different datasets addressing ***the same research question****" (Koricheva et al., 2013); (2) it implies that plots within datasets are not independently and identically distributed, which is not indicated in InsectChange; and (3) it introduces the problem of the "false baseline effect" (Didham et al., 2020), i.e., any nonrandom bias towards an above-average or a below-average starting point in a time series comparison, with a subsequent bias in the overall trend estimation. Therefore, because of these often artificial situations, which lead to below-average starting points much more frequently than above-average starting points, the insect trends obtained from InsectChange data (van Klink et al., 2020a) for freshwater and terrestrial realms are most likely overestimated.*
How could data selection be improved in InsectChange? First, to reach more robust and meaningful conclusions, the best way to proceed would be to select more homogeneous datasets enabling testing of a single clear hypothesis, or alternatively to control for heterogeneity among studies with statistical analyses that take these differences into account with predictor variables. For controlled experiments, it would be relevant to consider only control sites. For other datasets, care should be taken to ensure the representativeness of situations and drivers in terms of sites with or without disturbance and in terms of timings of disturbance, and disturbance types could be weighted according to their frequency (Cardinale et al., 2018). Maps of human impacts on ecosystems, for example, could guide the choice of data and/or their weighting (Gonzalez et al., 2016)".

R2: Figure 4g: I would use a different palette for this plot, as the same colors are also used in panels 4a to 4f but are associated with different meanings.

LG & MD: we have changed the palette of former Figure 4g (now Figure 5g) as suggested.

R2: Lines 187-188: not entirely clear to me what is meant by 'plots with distinct geographical coordinates'. Are these the plots that do not belong to groups of plots that were assigned a unique pair of coordinates? Please, clarify.

LG & MD: Our section 4.1 on the inadequacy of geographic coordinates now details cases where a pair of geographic coordinates was attributed to a single plot in InsectChange versus cases where several plots shared the same pair of geographic coordinates. It also summarises cases where these identical geographic coordinates are not problematic given the scale for local analysis, because sampling points were very close to each other, and cases where these identical geographic coordinates are problematic, because sampling points were actually too far from each other. The detailed information is now available in our supplementary file *CroplandCover.xlsx*. It is summarised in the new Figure 6 of the manuscript.

#————————

R2: Appendix S4 (no line numbering)

This appendix has been removed.

R2: Point #1 of the list of possible causes of the incorrect assessment of local cropland cover: the authors highlight a mismatch between the cropland cover reported in original studies and that estimated from remote sensing. I would not be surprised by this mismatch if in the original studies cropland cover was estimated at a truly local scale. ESA-CCI remote sensing data provide worldwide cropland cover at a (not that coarse) spatial resolution (300 m x 300 m) and note that InsectChange aggregates data across an even larger spatial window. I believe this sort of mismatches are inevitable when comparing (truly) local vs. global remote sensing data.

LG & MD: we have extensively reworked section 4 of the manuscript. We now state (p. 17, lines 486-493): "*The main reason for inadequate cropland cover assessments was the inaccurate interpretation of satellite images by the ESA-CCI database (CroplandCover.xlsx), notably because grasslands, heathlands, steppes, barrens, prairies, shrublands, marshlands, natural vegetation areas, parks or golf courses may inaccurately be coded as croplands (Peng et al., 2017; Liu et al., 2018), and the representation of land cover is imprecise when used at a local scale composed of nine 300 m × 300 m squares with rough cropland cover assigned to each of them (63.2% of inadequate cropland cover assessments, CroplandCover.xlsx, example in Figure 7b)*." As detailed in our supplementary file *CroplandCover*.xlsx, most often this issue is not linked to the 300 m x 300 m spatial resolution, it is just that large areas of grassland, prairie, park etc., are inadequately coded as croplands in ESA-CCI. Therefore, the issue is not mainly related to a question of truly local remote sensing data. Of course, mismatches are inevitable; we believe that our additional work to quantify minimum overestimations and underestimations of cropland cover (boxplots in Figure 7) shows that these mismatches are important enough to call into question the reliability of an analysis that would be based on the assessments of local cropland cover reported in InsectChange. We have now removed Appendix S4, which had become redundant with the main text of the manuscript and our supplementary file *CroplandCover.xlsx*.

R2: Also, I am wondering why the visual interpretation of satellite images (mentioned in the first sentence of point #1) should provide a better estimation of cropland cover than that provided by ESA-CCI, given the difficulty of distinguishing between grasslands and croplands acknowledged also by the authors.

LG & MD: we have clarified in the main text of the manuscript, section 4.2, that we combined the visual interpretation of Google Earth satellite images around the correct sampling areas with information available in the original studies, in other publications, on satellite images from Landsat 8 or Sentinel 2 for more dates, on the internet and in ESA CCI. As detailed in *CroplandCover.xlsx*, we only based our interpretation on clearly identifiable parts of the land cover, and most often we did not conclude on the basis of Google Earth alone; when we did, the interpretation was visually very clear. We have checked in detail our assessment, included additional sources of information and completed our comments.

R2: In point #1, I missed a more rigorous (and objective) quantification of the mismatch between locally estimated cropland cover and that derived from ESA-CCI (e.g., average difference between cropland cover reported in original studies or visually estimated from satellite images vs. that reported in InsectChange). This could help readers understanding what is real magnitude (and potential impact) of this issue on InsectChange.

LG & MD: we have extended our former analysis to include, when possible, an assessment of a minimum or maximum cropland cover per plot based on the measurement of clearly identifiable parts of local land covers. This now allows us to show a minimum error in InsectChange estimates for overestimated and underestimated cropland covers, as summarized in two boxplots in Figure 7 (see answer to one of your previous comments above).

R2: Finally, last sentence of point #1 sounds to me more as a critique to the ESA-CCI land cover product than to InsectChange.

LG & MD: In the main text of the manuscript we now write (lines 486-493 p. 17): "*The main reason for inadequate cropland cover assessments was the inaccurate interpretation of satellite images by the ESA-CCI database (CroplandCover.xlsx), notably because grasslands, heathlands, steppes, barrens, prairies, shrublands, marshlands, natural vegetation areas, parks or golf courses may inaccurately be coded as croplands (Peng et al., 2017; Liu et al., 2018), and the representation of land cover is imprecise when used at a local scale composed of nine 300 m × 300 m squares with rough cropland cover assigned to each of them (63.2% of inadequate cropland cover assessments, CroplandCover.xlsx, example in Figure 7b)*". The limitations of ESA-CCI are discussed in the articles quoted in this paragraph, and we believe that this database can perform well in some contexts, less in others. In our view, it is InsectChange methodology of using this database for their matching that is in question.

R2: Point #2: once again, it is not clear to me what the authors mean by 'inaccurate location'. Please clarify.

LG & MD: In the main text of the manuscript, section 4.1, we now clearly define that "*[t]he adequacy of these local-scale indicators hinges on the premise that, for each plot, the geographic coordinates assigned to the plot in InsectChange are precise enough to point to the insect sampling area, and that this sampling area is included in a 900 m × 900 m square (hereafter referred to as a "local-scale square", Figure 6a)*" (lines 394-398 p. 14) and that an inaccurate location means that InsectChange geographic coordinates were "*included in a local-scale square that was outside the sampling area*" (lines 444-445 p. 16).

R2: As already pointed out above, I would avoid reporting details about figures in the main text in an appendix (and vice-versa).

LG & MD: The details on former Figure 5b (now Figure 7b) that were previously provided at the end of the former Appendix S4 have been moved to the caption of Figure 7. We suppressed this Appendix S4 and provided the main results in the text and the details in our supplementary file *CroplandCover.xlsx*.

#————————

R2: Figure 5: I would include information to interpret Figure 5b in the caption. Concerning Figure 5a, what are the numbers on the y-axis about?

LG & MD: Information to interpret former Figure 5b (now Figure 7b) has been moved from the end of former Appendix S4 to the caption of Figure 7 (also see comment above). In Figure 7, the y-axis is the number of plots, as now indicated on this axis.

R2: Figure 5 reports a rather extreme case of mismatch between cropland cover reported in InsectChange and what estimated by the authors (or found in the original source). I think that it

would be good to show what is the average mismatch between the information derived from ESA-CCI and that estimated by the authors. This could be done using data reported in the CroplandCover Excel file. I believe such a comparison would more objectively demonstrate the magnitude of the issue.

LG & MD: As detailed in an answer to one of your previous comments, this information was actually not available in our former CroplandCover Excel file, but we have performed an important additional work to include this in that file and in the main text (see boxplots on former Figure 5, now Figure 7).

R2: Line 215: I would be more cautious here and state that errors, inconsistencies and methodological issues found in InsectChange could bias estimation of temporal trends in insect change (rather than claim they fully undermine what found by, e.g., van Klink et al. 2020). Indeed, although several concerning examples on the misuse of data included in InsectChange are highlighted in the comment, authors mostly focus on dataset-specific issues and do not quantify their overall impact on the conclusions of previous works based on InsectChange.

LG & MD: We agree that we did not quantify the overall impact of the issues we identified in InsectChange, and therefore that we cannot claim that they fully undermine the results obtained from this database. We have rephrased "*The numerous problems underlying InsectChange call for corrections and an extreme vigilance in its use. They fully undermine the results obtained so far from this database […]*" as: "*The numerous problems affecting the InsectChange database call for corrections and extreme vigilance in its use. They call into question the results obtained so far from this database […]*" (lines 529-532 p. 19). We agree that "fully undermine" was too strong but we believe that the content of our comment makes a sufficient case to show that the issues are not marginal, as "could bias" suggests.

#————————

R2: Appendix S1 (no line numbering)

Minor issue: I would move the "Presentation of the different appendices", which describes all appendices, outside Appendix S1.

LG & MD: If possible, we would prefer to keep this description here in order to avoid creating a new supplementary file.

R2: I would move the section about "Other problems not included in our analysis" at the end of Appendix S1.

LG & MD: If possible, we would prefer to keep this section here as we are afraid that it may not be visible at the end of the long Appendix S1.

R2: Study 70*, line 15: shouldn't be 'beyond 1996' here?

LG & MD: Thanks, this has been corrected.

R2: Study 300: a reference for ESA-CCI is missing.

LG & MD: Thanks, the reference has been added.

R2: Study 375*, last sentence of the paragraph: this statement is vague as no measure of the difference between cropland cover from ESA-CCI and that visually estimated from satellite images is reported.

LG & MD: This last sentence in the comment of Study 375 was: "*Finally, the estimation of the local cropland cover provided in InsectChange is not consistent with satellite images*", we have now added: "for Plots 243 and 259 and unclear for Plot 256".

The paragraph on the first page of Appendix S1 specifies: "*Details on our assessment of the adequacy for a study at a local scale of the geographic coordinates provided by InsectChange and/or the adequacy of the local cropland cover provided for terrestrial plots are presented in CroplandCover.xlsx*", because we cannot copy all the details provided in *CroplandCover.xlsx* in Appendix S1.

In *CroplandCover.xlsx*, the comments for these plots are the following:

- Plot 243: "*The local cover cropland was 0% (and not 17%) on the basis of satellite images in 2014, including 2014 Landsat 8 images. In ESA-CCI, the two cells to the north and to the northwest of the one with the plot were coded as mosaic cropland (>50%)/natural vegetation (tree, shrub, herbaceous cover) (<50%) (ESA-CCI code 30) while they are forest areas.*"

- Plot 259: "*In 2014, the local cover cropland is 0% (and not 22%) on the basis of satellite images (2014), including Landsat 8 images in 2014. In ESA-CCI, the cell with the plot and the one to the west of the plot are coded as cropland (ESA-CCI code 10) while they include a parking lot and a forest area.*"

- Plot 256: "*Satellite images in 2014 show an area that may be a crop plot northwest of the site.*"

In our opinion, these comments provide enough details to understand the reason for our assessment.

#————————

We hope that we have successfully addressed your concerns. Thank you for your helpful comments, which, we believe, have enabled us to improve the quality of our paper.