**RESPONSE TO REVIEWERS**

**GEB-2023-0186**

**Functional convergence underground? The scale-dependency of community assembly processes in European cave spiders**

EDITOR'S COMMENTS TO AUTHORS  
Editor: Gonzalez, Angelica  
Comments to the Author:

Dear Dr. Mammola,  
Thank you very much for submitting your manuscript "Functional convergence underground? The scale-dependency of community assembly processes in European cave spiders” (GRB-2023-0186) to Global Ecology and Biogeography. We appreciate the work you have accomplished, but there were substantial concerns about the paper. I will not be able to accept this manuscript for publication in GEB in the current state.

First, I would like to apologize for the delay in getting this decision to you. I spent several months out of the country performing field work, and I had no or very limited internet during that time. I have obtained two reviewers with extensive expertise in functional traits, functional trait-based approaches, community assembly, and arthropod ecology. The reviewers and I agree on the promise of this study in analyzing community assembly processes focusing on an understudied but interesting group, cave spiders. The reviewers also acknowledge the quality and scope of the dataset used in this study. However, the story developed in the manuscript, in its current form, raised major and fundamental criticisms to your study.

Both reviewers questioned the depth and details of the information provided to fully understand the study. In my own assessment, I concur with reviewer 2 that the despite your questions and analyses are interesting and well performed, these do not seem to fully accomplish the goal of disentangling the relative importance of environmental filtering and limiting similarity. You may not have a final answer, but we agree that we can still obtain insights into those processes. Both reviewers had a difficult time to follow the methods used in this study, among these the lack of clarity in the trait selection criteria (see also advice on additional information related to the traits included in the study), its potential influence on the multidimensional functional space of this group of spiders, and some more discussion on the potential biases of the method of choice (hypervolume). In addition, both reviewers had some problems following some of the statistical models used to test for changes in beta diversity, null models, etc. I think this is due to the lack of details along the method section, which needs to be improved. Another major issue, is the too succinct discussion of the relevance of both processes environmental filtering and limiting similarity, as highlighted by reviewer 2, both processes can be acting in tandem. Your interpretation that under dispersion as environmental filtering and overdispersion as limiting similarity needs more nuance, as pointed out by earlier work, most null models can identify non-random patterns of species associations, but are unable to distinguish the relative effects of limiting similarity and environmental filtering as the mechanisms underlying the non-randomness (see Gotelli & Ulrich 2012 Oikos, D’Amen et al. 2017 Ecography, among others).  
  
There are also some criticisms on the clear development of your hypotheses, which need to be revised. In addition, the discussion did not provide enough concrete comparisons and contrasts when interpreting your results. Finally, as highlighted by reviewer 1, the abstract needs more nuance to better reflect the main findings of this study.

Because we all see potential on this study, I encourage you to revise and resubmit your manuscript to Global Ecology and Biogeography. However, all constructive comments provided by the two reviewers must be taken into consideration into preparing a revised version of this manuscript.  
  
Dr. Angelica Gonzalez, Editor

**RESPONSE:**

**Dear Editor,**

**Thanks for handling our submission and the useful guidance on how to best address the reviewers comments. We believe the critical feedback by the reviewers helped us to improve different aspects of the manuscript.**

**To summarize, in this revision we have:**

**1) Recast the hypothesis framework following suggestions by ref. 1 and 2;**

**2) Better explained the choice of traits and their nature in the main text (this info was previously only available in the Supplementary materials);**

**3) More clearly explained methodological choices and details (hypervolumes, null mdoelling, BBGDM);**

**4) Recast the abstract to make it more focused on the numerical results;**

**5) Better discuss the trait space of spiders (PCoA), including adding a new figure summarizing the shape of the trait space;**

**6) Incorporated all other minor suggestions by the two referees.**

**Please find full information on the nature of these revisions in the point-by-point responses below. Main change to the manuscript are also highlighted in yellow.**

**We believe the manuscript greatly improved as a result of this constructive round of revision, and we are looking forward to hearing your opinion.**

----------------

REVIEWER COMMENTS TO AUTHOR  
  
Referee: 1  
  
Comments to the Author  
The study addresses alpha and beta functional diversity of European subterranean spiders and showed a predominant trait clustering, which at southern latitudes with more extensive karst areas, more variable caves and higher temperature ranges turns into predominant trait overdispersal. On the other hand, patterns of beta diversity indicate species turnover while trait space is kept saturated, that several environmental factors prompted beta diversity and that overdispersal is detected on a local level. The study has a good dataset, sound methodology (which I am not familiar with entirely), relevant questions and interesting conclusions. Doubtless, it is a study to be published by Global Ecology and Biodiversity, subject to revisions. I have four major concerns and several minor included directly into the annotated manuscript.

**RESPONSE: We thank the reviewer for investing time in reading and commenting our manuscript, the positive attitude towards it, and all the constructive feedbacks.**

Firstly, I had problems to understand the part of the manuscript related to beta diversities. The first part of work, devoted to alpha diversity, is straightforward. However, I am not familiar with models used in modelling beta diversities and I had hard times to understand what authors did, how they handled the data and how to read results. There are multiple beta diversities, and diverse SES values modelled against the gradients. It is hard to follow which SES values authors refer to. If the wordcount does not allow, I strongly suggest that authors in a greater detail explain the procedure in Supplement.

**RESPONSE: In our study, we used generalized dissimilarity modeling (GDM) to explore the relationship between dissimilarities and environmental as well as spatial gradients. For beta diversity analysis, we applied a null model approach similar to alpha diversity, involving the shuffling of species traits to derive null functional spaces. These null functional spaces were then compared to generate null beta diversity values. While traditional approaches contrast these null beta diversity values with observed beta diversity values to compute Standardized Effect Size (SES), our focus was on discerning whether the observed pattern was stronger or weaker than expected along the gradient. Particularly, we aimed to investigate whether the likelihood of stronger or weaker responses than expected varies along the environmental gradients. Notably, deriving SES values for beta diversity is challenging due to the absence of distributional families for SES values in the context of GDM, which only applies a binomial family of error distribution bounded between 0 and 1. To address this, we used null beta diversity values to build a GDM model, extracting fitted curves for all predictors using the null beta diversity as a response variable for multiple randomized sets (n=999). The key modification is in deriving SES values from these fitted curves rather than directly from the beta diversity matrix. By constructing a GDM model using null values, we generated fitted curves and compared them with the actual fitted curve to estimate SES values for the fitted response. This approach enables us to assess whether the fitted response of dissimilarity is stronger or weaker than expected given changes in species composition. We believe this comprehensive explanation provides a clearer understanding of our methodology and welcome further discussions or additional clarifications in the supplementary material if necessary.**

Additionally, an important part of conclusions is confronting present results with results of the past study. If only possible (with respect to journal requirements), I suggest to reproduce the past results once again, to make study and discussion easier to reach. I acknowledge this first problem I have may be simply due to my limited knowledge and this comment may turn to be irrelevant.  I therefore ask editor for additional evaluation.

**RESPONSE: We have now modified the text and removed the mentioning to the other study at least in the way it was done. However, we believe that replicating results would not be an optimal practice as it could qualify even as self-plagiarism. ch**

My second concern is related to hypothesis H1b, and is in contrast to results. Why limiting similarity would take the role in ecologically constrained space / dispersal? Limiting similarity is a concept intimately linked to ecological disparity. If there is no ecological space, disparity – to my knowledge cannot evolve (unless we speak of temporal differentiation, e.g. different phenologies). The only effect in this case would be lower species number, as competitive interactions exclude one of the species, but this is not tested. I am missing some logical explanation; please, elaborate.

**RESPONSE: Thanks. Your comment and those by referee 2 made us realize that the hypothesis framework was not clear. We now recast the whole hypothesis section, as follow:**

**ADD TEXT**  
Thirdly, the Discussion is sometimes too abstract. I labelled some points, where authors very technically interpret the results, but I think the manuscript would benefit if they tried to explain results more in depth.

**RESPONSE: Thanks. We revised the discussion following your suggestions on the annotated file.**

Fourth, abstract should be rewritten and made more explicit. I think in present form it may not attract the potential readers.

**RESPONSE: Thanks. We revised the abstract following the specific suggestions on the annotated file.**

Overall, it is an important piece of work, but I think it would benefit if it was better explained.  
**RESPONSE: Thank you.**

------------------------------------------------

Referee: 2  
  
Comments to the Author  
Functional convergence underground? The scale-dependency of community assembly processes in European cave spiders  
  
  
GENERAL COMMENTS  
In the present study, the authors tackle an important question in community ecology regarding the relative contribution of environmental filtering and limiting similarity in driving the functional assembly of communities. To this end, the authors use a neat study model in the form of subterranean spider communities, which lives in caves distributed throughout Europe. They leverage an impressive and already published database on cave spider distribution, and functional traits.  
  
They first found that most subterranean spider communities displayed trait dispersion that did not depart significantly from random. Yet, overdispersed communities tended to be positively correlated to annual temperature ranges, karst area and negative cave drop. Last, the cave development was more strongly related to changes in functional beta-diversity than did the geographic distance between any two caves. Overall, the manuscript is very well-written, and despite many specific comments required to address some lacks of clarification in the main text I have two main comments.

**RESPONSE: We thank the reviewer for investing time in reading and commenting our manuscript, the positive attitude towards it, and for the useful constructive feedback.**  
  
First, we understand in the M&M that this study uses only a subset of all available traits, the selection criteria are not provided and we can wonder how those choices can influence all the results (PCA, hypervolumes, null modeling and SES). To me one of the great novelty of this study is to explore the multidimensional functional space of cave spiders. Therefore, a deeper look into this space would strengthen the study (see also my specific comment to display and discuss the PCA).

**RESPONSE: We now clarified methodological aspects and added a new figure illustrating the trait space (see answer below; important note: we didn’t ran a PCA, but a PCoA). Furthermore, we provided more details on the selected traits and undelrying reasons (see answer below).**

In addition, the method of hypervolumes is used because they are ‘less sensitive to outliers than convex hulls’ then I wonder to what extent this hypervolume method properly estimates the dispersion (ie including outliers) of communities upon which the entire study is based on. How are outliers defined here?

**RESPONSE: As specified in the M&M, the main reason we decided to use hypervolumes is that ‘*they allow the detection of areas of higher or lower density in the trait space, thus representing uneven probabilities of finding a species with a given trait combination throughout the boundaries of the trait space*’. We mention outliers as an additional reason because, given the uniqueness of some traits of cave species, we can expect several communities to have functional outliers (and a convex hull would overestimate functional richness in such cases). We now removed the mention to outliers, to avoid any confusion.**

**On a side note to answer your question: hypervolumes estimate dispersions as the average distance between a sample of stochastic points defining the hypervolume (not the real observation) and the hypervolume centroid. However, we believe this is not really relevant to mention here as we focused on the Richness, rather than the Dispersion component of FD.**  
  
Second, I think that the authors haven’t properly dissected the relative influence of environmental filtering vs limiting similarity. They interpret underdispersion as environmental filtering and overdispersion as limiting similarity. Therefore, a given subterranean spider community is either under- or overdispersed, while theoretically both processes act simultaneously on any given communities (i.e. environmental filtering can result in overdispersed communities).

**RESPONSE: We fully agree with the reviewer. In fact, we have acknowledged this in the discussion (e.g., L391+ of the initial submission “Whereas the direction of SES for functional richness was predominantly towards underdispersion (Figure 1b), the majority of values were close to zero, meaning environmental filtering and limiting similarity were both acting in equally weak or strong, but opposing, directions.”). We now gave more prominence to this view, by mentioning this as early as the Abstract.**  
  
If I understood it well, the null modeling approach presented here mainly test for the non-random assembly of species whether it is due to environmental filtering and/or limiting similarity. In addition, an hypervolume does not allows evaluating the differences between species inside the volume (i.e., limiting similarity), because the same hypervolume can be generated by a different number of species depending on how dissimilar they are.  
  
Therefore, to further understand the relative importance of environmental filtering and limiting similarity, I would suggest to maintain environmental filtering constant while varying limiting similarity. A first step toward this goal would require a better understanding of the relationships between hypervolumes, SES and species richness along environmental filters.

**RESPONSE: We agree with the reviewer, we tried to add more nuance in the discussion of the patterns observed in our study. Hypervolumes, in itself, are not able to distinguish between environmental filtering or limiting dissimilarity. However, our analysis allow understanding whether changes in community composition have a stronger or weaker selection in the trait component of communities. For example, if we observe SES values significantly small (< 1.96) we can understand that community composition, changed in a way that different species between communities are more similar in traits than would have been expected if the species assembly was random. This may indicate (if we selected relevant traits) that the environmental forcing in those communities is strong enough to not allow the presence of different traits. In the opposite case, we can speculate that there is enough niche available for other traits to co-occur be it because of weak environmental filtering or limiting similarity. We agree that patterns are not exclusively generated by competitive exclusion, but also niche availability and we have therefore modified our discussion to fit the explanation within this point of view.**

SPECIFIC COMMENTS  
  
ABSTRACT  
L22: independent of

**RESPONSE: Done**  
  
L27: the result section is too vague, the first sentence is not a result of this study but rather a definition of environmental filtering and limiting similarity. Please give more details about what scale and environmental gradient? In addition, I think that you forgot to mention that one of your main results is that 97.8% (100- (7underdispersed + 1overdispersed)/367\*100) of the subterranean spider communities display trait dispersion that do not depart significantly from random. This result is also barely discussed to better understand why this is the case.

**RESPONSE: Thanks for these suggestions. We recast the ‘result section’ of the Abstract as follow:**

**ADD TEXT**

L35 after reading the entire manuscript, first I am not sure to figure out how your study has reconciled the relative importance of environmental filtering and limiting similarity and second I am not sure what particular ‘conceptual foundation’ this study brings to the study of community assembly. Please clarify these points using your main findings.

**RESPONSE: yes, this was probably an oversterach from our side. We recast the conclusion as well to be more focuse on the actual findings. Now it reads as follow:**

**ADD TEXT**  
  
  
INTRODUCTION  
L77-85 I would consider the literature about tank-bromeliads (doi:10.1016/j.tree.2004.04.010) for insightful examples of island-like labs.

**RESPONSE: Thanks, we now cited plant-island examplee and this useful reference.**  
  
L81 remove the ‘]’

**RESPONSE: done.**  
  
L109 this second ii) test is not clear. The first makes sense, as it is well described in the introduction but not sure about the second one. Perhaps some information is missing here. Why would you expect functional diversity to decay along environmental gradients? I guess it would depend on the identity of the environmental gradients (please name them) and on the intensity and direction of variation of these gradients. Please clarify. In addition, this second point is slightly different from the one present in the abstract L11.

**RESPONSE: Good point. Here, the implicit modeling assumption is that the similarity of communities can only 'decay,' implying that beta diversity always increases along gradients (see https://onlinelibrary.wiley.com/doi/full/10.1111/geb.13513). However, we agree it was not entirely clear in the text, so we have rephrased it to a more neutral version ('how the relative contribution of these driving forces changes along environmental gradients'), which is also consistent with the abstract.**  
L116 unless I misunderstood the link, I think that H2a (functional vs taxonomic turnover) is disconnected from its explanation (stringent conditions should reduce the volume of trait space regardless of spatial scales). In other words why do you expect functional turnover to be lower than taxonomic turnover. To me there are two interrelated reasons. First, there is functional redundancy. Second, this functional redundancy (lower functional turnover) may be a mathematical/methodological artefact of describing many species on a few trait axes. In other words, what if you would integrate more traits in the analysis.  
In addition, to me the explanation of H2a (stringent conditions should reduce the volume of trait space regardless of spatial scales; #1) is more related to H2b, which already has an explanation (niche availability; #2), so this explanation #1 is somewhat redundant with that (#2) of H2b. Please clarify this.

**RESPONSE: This is a very good suggestion for simplifying the hypothesis framework. We now merged H2a and H2b into a single hypothesis, as follow:**

**ADD TEXT**  
  
MATERIALS & METHODS  
L126 in this ‘community-level data’ section it would be nice to give a bit more details on spider communities, simple descriptors like how many species in total and in average in each cave. The naive reader (me) may not want to go through all references cited to find this basic information.

**RESPONSE: good point. We included this sentence:**

**“The database includes 326 species [average (±s.d.) number of species per cave of 4.3 (±2.35); range: 0–15].”**  
  
L137 I would rather rename this section ‘Environmental and geographic gradients’ and introduce the geographic distance in this section (tested in the BBDGM section from L255)

**RESPONSE: Done as suggested.**  
  
L140 I believe this is Mammola et al. 2019a, b you are referring to?

**RESPONSE: Yes. Corrected.**  
  
L159 ‘proxy’ is not italicized while it is L162, pick one throughout

**RESPONSE: Done.**  
  
L166 in this ‘functional traits’ section I’d like to have the number of traits, as well as the number of traits modalities for each of the three cited categories (morphology, adaptation, hunting strategies). By the way, very nice to see a figure (Fig1) with the distribution of numerical values for each trait.

**RESPONSE: All this info was available in supplementary material Table S1, but we agree a summary in the main text is useful, and we now added this information. Trait modalities are avilable in Figure 1, which is already in the main text, so we avoided repeating the info in the text.**  
L169 ‘we selected a subset of traits’ it would be nice to present the selection criteria and the number of traits retained compared to the total number of traits. See my general comment about a discussion on how integrating the entire trait database would change the results of the present study.

**RESPONSE: The selection of traits and their meaning was fully explained in the paper illutstrating the trait database (Mammola et al., 2022 Scientific Data). Yet, we agree providing more information in the main text will ease the fruition of this specific manuscript. Therefore, we expanded section “Functional traits” as follows:**

**ADD TEXT**L185 remove ‘refs.’ and add the ref for 2009 before that of Palacio

**RESPONSE: Done.**  
  
L201 ‘we excluded caves’ instead of ‘we filtered caves’

**RESPONSE: Done.**  
  
L206 for the reader non-familiar with the ‘de Bello’ approach would you please explain the essence of it in a few sentences, because it seems an important choice in the overall analysis.

**RESPONSE: We now provided details about the method, as follow:**

**ADD TEXT**

L233 not clear how the null distribution is obtained. First, I guess that species are in rows and traits in columns. Second, what do you mean by ‘randomized’? Do you mean randomly permuted without replacement? In other words, for instance with a cave with three species A, B and C (with an observed hypervolume Vobs), at each iteration ‘i’ a random set of three species (containing or not A, B or C) is used to calculate a random hypervolume Vi? If this is the case then the 8 significant spider communities (7 underdispersed and 1 overdispersed) may be due to particular sets of species adapted to particular environmental conditions (temperature, kart area, and negative drop). In this sense, overdispersed communities may be the result of relaxed or less intense environmental filtering than increased limiting similarity.

**RESPONSE: We agree with the referee. We have fixed the text accordingly to explain the randomization and we now avoid the terminology “limiting similarity” throughout the manuscript.**   
  
L235 ‘(see next section)’, in fact diversity measures are explained in the previous section, so I would remove this statement between parentheses

**RESPONSE: Done**  
  
L235 how do you estimate/define ‘the deviation of observed values from the null distribution (eg mean (abs(null SES) >= abs(observed SES))?

**RESPONSE: CAIO**  
  
L238 not sure if I understand this test: ‘estimate non-parametric effect sizes using probit transformed p-values’ to me you estimate the significance of SES using probit-transformed p-values (obs vs. null distribution) not the SES directly. Or you don’t estimate the significance (only relative SES) but then at what threshold do you say it departs ‘significantly’ from random? In the legend of Figure 3 we understand that significant effects are Rank < 0.025 | > 0.975, which I agree with, but if that’s the case then it should be clearly stated in the M&M.

**RESPONSE: We appreciate the opportunity to provide further clarification on the test involving probit-transformed p-values for estimating non-parametric effect sizes (SES). We have rigorously followed the Lhotsky et al. (2016) guidelines, where the authors explicitly mention that to measure the strength of convergence/divergence, they have built upon the approach proposed by Chase et al. (2011) and Bernard-Verdier et al. (2012), who utilized linearly transformed p-values as an effect size (ES) by replacing the linear transformation with a probit transformation. Therefore, in our study, SES was indeed estimated as probit-transformed p-values. The p-values were calculated by determining the proportion of instances where the estimated values were less than the observed values, plus half of the instances where the estimated values were equal to the observed values, divided by the total number of estimated values. In R code, this calculation is expressed as:**

**p\_value <- (sum(estimated < observed) + sum(estimated == observed)/2) / length(estimated)**

**This calculation was then subjected to a probit transformation to obtain the SES. Notably, for our determination of significance, we applied a threshold of Rank < 0.025 or > 0.975. We trust that this explanation clarifies our methodology, and we are open to further discussions or additional clarifications as needed.**

L249 To me, including spatial autocorrelation in models is useless because it is nothing but the residual variance due to unmeasured environmental predictor, so I would remove it from analyses. It makes only sense when used as predictor and as a proxy for dispersal limitations with explicit tests of underlying processes (linear, quadratic, exponential). I know this comment is reviewer-dependent so I would at least indicate whether the results change when including it or not.

**RESPONSE: Thank you for your insightful comment regarding the inclusion of spatial autocorrelation in our models. We appreciate your perspective and wish to clarify that our decision to incorporate spatial autocorrelation is driven by the goal of addressing the skewness of residuals associated with spatial structure in environmental variables, rather than directly measuring the effect of spatial component. Our models indeed did not improve by including spatial autocorrelation, but as the referee well said, inclusion or not of such component is highly reviewer dependent. We clarified now in the results section this detail.**

L251 although common in ecology, but really not necessary is to standardize predictors in regression analyses, it makes the interpretation far more difficult.

**RESPONSE: Beyond the reason given (easier comparability), it was also needed to ease model convergence. We now mentioned this.**  
  
L274 ‘SES of a predictor’ not sure to understand here I thought it was the SES of a given alpha- or beta-diversity measure (not that of a predictor or environmental filter)?

**RESPONSE: Good catch. This was a wrong wording. We now corrected the sentence.**  
  
L274 we understand here that a ‘geographic gradient’ is tested as a predictor but it is not mentioned in the M&M (see my comment above about the ‘Environmental and geographic gradients’). In addition, L250 ‘an exponential correlation structure’ is used so I wonder whether a similar relationship (exponential) is used for the modelling of the geographic gradients at the beta-diversity scale?

**RESPONSE: We now mention the geographic gradient variable in the section “Environmental and geographic gradients” (see response to your previous suggestion). There’s no need to superimpose a exponential correlation structure in BBGDM. The model explore the relationship with the grographic distance and can assume different nonlinear shapes (although in a monotonic fashion, due to the mathematical nature of beta diversity). This is explained in the M&M, at Line XXX.**  
  
RESULTS  
L281 as a first paragraph in the result section, I think that it would be neat to have a look at the main PCA that was used to estimate all hypervolumes, and to display which traits were correlated to each of the three PCA axes (see Céréghino et al. 2018: DOI: 10.1111/1365-2435.13141 for an example). I wonder if the PCA would change following the de Bello approach or not. Would it be possible to run the PCA analysis with weighted and unweighted traits?

**RESPONSE: We have now included a PCoA plot in the main text (Fig. XX) and a table with the contribution of the the traits to the three axes in the supplement material (Table. SXX). We had previously checked about the trait distribution using both the unweighted and weighted approach and found that results did not change. We choose the weighted approach because we agree with the rationale of the data treatment.**   
  
L294 ‘Patterns … Figure 3)’ this sentence is uninformative, at least say how many environmental gradients.

**RESPONSE: We now specified the gradients.**  
  
  
DISCUSSION  
L319 baseline for what ? To me a baseline is related to an ecosystem state before a given disturbance.

**RESPONSE: True, rephrased.**  
  
L354 what is this number (51) ? If it is a reference please add it.

**RESPONSE: Sorry, this was a typo derviing from a previous formatting of references. Now we have corrected it.**  
  
L361 not sure to follow where you see an inflection point at 10 m? not clear by looking at Figure 3

**RESPONSE: Thanks for spotting this. The correct value is indeed 5m.**

L368 the difference in responses between taxonomy and functional traits is also an artefact of using more species than functional traits. Especially because in this study only a subset of the trait database is used.

**RESPONSE: Thanks for this important comment. We aren’t fully convinced with this interpretation. Different traits combinations for subterranean spiders have been tested to a great length in previous analyses (e.g., Mammola et al., 2022). Given the redundancy in some traits, this selection of trait gets very close to maximizing the information contained in the trait database (see previous answer and additions to the text, explaining the choices of specific traits), while minimizing the number of missing data and hence the need for trait imputation. We now mentioned this is the M&M section.**

FIGURES  
  
L640 remove redundant information ‘after excluding . . . < 3’ (also present in M&M)

**RESPONSE: Done.**  
  
L652 Figure 3 : I am a bit surprised by the ranges of variation of observed values in gradients of panel a) and c). For instance, the elevation varies from 3 to 7 m asl, the geographic distance varies from 0 to 40 km, which both sound extremely surprising when looking at the map in Figure 2. In addition, the LGM ice distance and the precipitation values seem weird. Even log-transformed values don’t make sense. According to the map in figure 2, two distant caves (Norway – Turkey) would be roughly 3500 km apart (as the crow flies), so a log(3500) is 8.2 not even close to 40. Please clarify those observed gradients.

**RESPONSE: We apologize, there was a mistake in the unit of the variable. We wrote the geographic distance is in [km], but it was actually decimal degrees (coordinates are expressed in WGS84 decimal degrees). Indeed, the two distant-most caves are ca. 40° apart, and so the range of the variable is 0–40°.**

**CHANGED**

L661 how is ‘expected given taxonomic composition’ estimated?

**RESPONSE: With the same null modelling procedure used before. We now changed “taxonomic composition” to “species richness” to make this more apparent.**