



Reply on RC2

Laura Anahí Macario-González et al.

Author comment on "Geodiversity primarily shapes large-scale limnology and aquatic species distribution in the northern Neotropics" by Laura Anahí Macario-González et al., Biogeosciences Discuss., <https://doi.org/10.5194/bg-2021-298-AC1>, 2022

Reviewer 2.- We appreciate all the comments and suggestions, they will significantly improve the manuscript. Below we provide answers to all comments. We have tried to address most of the suggestions made by reviewer 2, but in case this is not possible, we provide a detailed explanation.

Comment: Because the authors are evidently not native speakers of English, there are considerable numbers of grammatical mistakes and poorly-phrased passages, most of which are detailed below with suggestions for improvement.

Answer: Sentence structure and grammatical errors will be corrected by a professional English editing service or a native English speaker with appropriate scientific background.

Comment: In several places the text is over-generalized, for example Line 35 which refers to "aquatic biological composition" – in fact, this study is only about ostracods.

Answer: We agree and will modify the text accordingly to prevent over-generalization. We now refer only to freshwater ostracods throughout the manuscript and exclude the term aquatic biological composition. Ostracodes, however, are recognized as bioindicators of aquatic ecosystems and ecological interactions change. The group belongs to the base of trophic chain and changes on its composition is evidence of aquatic biological composition alteration. The same apply on interpretations of past environments, because ostracodes have abundant fossil record and they reveal changes in the biological composition in response of external variables. We will also modify the title.

Comment: One area that requires clarification is the enumeration of ostracods (Line 204). Apparently the investigators counted ostracods, but the details of counting aren't presented. How many per sample?

Answer: For this study, we counted specimens in a standard sample volume of 50ml, obtained from filtering 200 L of water in the littoral zones. Then, we used relative abundance for statistical analyses. All this information will be described in detail in the manuscript. Count data expressed as relative abundance is provided in supplementary material <https://doi.pangaea.de/10.1594/PANGAEA.940254>

Comment: It's also not clear what they mean by "identified three adults" passage.

Answer: We used three well-preserved adult specimens for microdissection and an accurate identification down to species level. As the target specimens are microcrustaceans, identification based on a single specimen can be misleading, identification based on at least three specimens are considered reliable. We will clarify this sentence in the manuscript.

Comment: Analysis seems to be based on presence/absence, but if count data are available for each sample, then ordination can be based on square root-transformed percentage data.

Answer: Species assemblages were analyzed based on presence/absence data for the NMDS analysis. Because species composition from different sites (latitudes) are evaluated and most species are not shared between lakes, values of species richness and abundances can be highly variable between lakes (highly dissimilar). The NMDS outcome largely depends on the similarity or distance-base index used, therefore, literature recommend using presence/absence data in cases when dissimilarity is relevant. Please see *Ecology*, 84(3), 2003, pp. 777–790 by the Ecological Society of America. For the CCA, both presence/absence and relative abundance data resulted in very similar ordination. In the first version of the manuscript, we present the CCA graph of presence/absence data to be homogeneous with NMDS analysis (both with Presence/absence data), but following the view of the reviewer, we will present now the CCA graph based on count data.

Comment: A more serious shortcoming is data handling. Apparently the investigators did not test each variable for normality. Skewed data should be transformed to produce a more normal distribution prior to ordination. For a clear methodology I recommend the most-recent edition of Tabachnick and Fidell's book *Using multivariate statistics*.

Answer: Environmental data, given different units of quantification, were standardized prior to analysis by subtracting the mean value and dividing by standard deviation. Then, normality was assessed for each variable. Ordination and statistical analysis were therefore performed in a pre-processed database with data normal distributed. As this process may be unclear in the manuscript, we will describe it in detail. We appreciate the recommended literature.

Comment: The authors should be careful in discussing diversity, which is not the same as richness. Also, such comparisons are difficult to interpret because collection size varies among the studies that they cite. Using 'alpha diversity' and 'beta diversity' would help.

Answer: We used Alpha diversity for SEM analysis and following the view of reviewer we will use it to compare our results with other studies conducted worldwide.

Comment: I suspect that the authors are correct, that geology (local bedrock, karst vs volcanic) is the main influence on ostracod distribution. It's not clear why they use the term 'geodiversity' when bedrock geology alone seems to be the main driver. Data on 'geodiversity', i.e. local diversity of geology (geology, geomorphology, hydrology) aren't presented for each of the lakes.

Answer: The influence of geodiversity on ostracode species distribution is most likely occurring at different hierarchical levels. For instance, we did not find evidence (in SEM analysis) that local bedrock alone explains the species distribution. Currently, the SEM analysis is under review, and we will include local bedrock as an explanatory variable as well. The results of this hypothesis will be discussed in the manuscript. We use the term “geodiversity” because we are providing site characteristics, such as elevation and bedrock type, which constitute part of the geodiversity. We do not provide detailed data on geomorphology and hydrology because this is beyond the scope of the study. We will evaluate the relevance the term “geodiversity” for the manuscript and based on the outcome of SEM analysis, we will define the main driver of limnology and species composition. Then, the title will be change accordingly.

Comment: Another driver seems to be understated, namely precipitation rates. Region YG is low rainfall, Region GSHN is generally higher rainfall. This correlates with karst/low elevation vs volcanic/highlands, but isn't rejected as a main driver. In order to conclude firmly that it's geology rather than rainfall, the authors should add another variable to their data set (i.e. mean annual rainfall at each lake).

Answer: The influences of annual precipitation, precipitation seasonality, air temperature seasonality on species composition was also evaluated. This data was extracted from the Worldclim data base. Precipitation variables were, however, highly correlated with major ions and cations, particularly Mg and Ca. We decided to use the latter for further statistical analysis, as their variability in lakes may explain more clearly and directly ostracode distributions. We will provide a detailed description of variable correlation in the new version of the supplementary material.

Specific comments –

Answer: We appreciate all these comments and have started to revise the manuscript.