Referee: 1

Comments to the Author

In this manuscript, the authors test the hypothesis that there is a latitudinal gradient in forest species. The specifically focus on terrestrial vertebrates (mammals, amphibians, birds, and reptiles). They use species distribution map shape files, data on species habitat affiliations, and forest layer shape files to delineate the proportion of species that are forest specialist and the forest characteristics of 96 x 96 km pixels across the globe. Overall, I think this is an interesting question, but it would be more interesting to readers of GEB if the authors looked at other habitat types in addition to forest to see if this is a general phenomenon for all habitat types (i.e., there are more habitat specialists in the tropics for all habitats) or if forests are unique in that respect, as the authors suggest. The authors clearly have the pipeline to conduct the analysis and it would enrich the paper greatly.

**We agree that looking at species associated with other habitat types would be very interesting. Whilst we have the pipeline to calculate the proportion of species associated with different habitat types (response variable), we would not have the same explanatory variables available as we have for the forest species. Namely, we are unlikely to have data on habitat loss as we do for deforestation, as well as data on the long-term presence of said habitat types. This would mean that model structure would differ between habitat types which we believe would complicate results. As such we believe analysis looking at the distribution of species associated with all habitat types is better suited to a separate manuscript, and have not carried out this analysis within this paper. We have however discussed whether we would expect the observed pattern to hold across species associated with different habitat types on lines 378 - 384.**

I also believe they authors could do a better job contextualizing their question/study in what we know about the LDG. While they cover, some hypothesis loosely, the authors leave out other hypotheses.

**Thank you for bringing this to our attention. We have focused our choice on hypotheses as those which we believe could most affect the proportion of forest species in an area.**

**We agree that we could add further hypotheses, though as there are more than 100 suggested mechanisms that cause the LDG, we cannot include all of them in this study. We focused on adding two mechanisms which have previously been shown to impact habitat specialisation, and therefore likely the proportion of specialist species. These are ‘niche packing’ as tested by the total species richness in an area, and extreme environments/dispersal limitation as tested using altitude. The first was added to highlight the differences between species richness and proportion of forest species, and the second was added based on the reviewer's suggestion of Janzen’s mountains are higher in the tropics hypothesis.**

**With the addition of these two mechanisms, as well as further additions to the introduction suggested by reviewers, we have added a table (Table 1) to our introduction which includes all hypotheses, as well as brief rationale behind each, and our expectations. Furthermore, we have added relevant LBG analogues to this table, thus contextualising our question more within the LBG literature.**

I really feel the intro can be tightened up to lead the reader into the study in a more focused way.

**We agree that the introduction could be tightened up, and have done this in two ways. Firstly, we have removed much of the text in the introduction, and shown this in a shorter, easier-to-read format within Table 1. In Table 1 readers can easily see our hypotheses, our expectations, and the rationale, without the longer prose that were previously used. Secondly, we have completely rewritten the first two paragraphs (lines 42 - 75) to highlight what our question is, how it is different to the LDG, and its importance.**

Further the details of the analysis/model are lacking in the text and need to be better described.

**We agree that more detail could be added about the model, we have almost entirely rewritten the model description in the methods section, to both make the model clearer and to describe our new analysis. This can be seen on lines 228 - 232, and 239 - 245. We have also added the full structure of the model to the supplementary as pseudo-code in the Model Structure section of the supplementary.**

I also think the modeling approach is questionable. I detail my comments and concerns below.

**We believe the original bootstrapping approach is appropriate, however we agree that using one single model would be easier for model interpretation and reproducibility. As such we use a single model approach per taxon using smoothed Gaussian processes to account for spatial autocorrelation in the “bam” function from the mgcv package in R. Full details of this can be seen on lines 239 - 258.**

1) GIS/MAPPING

I do not have a lot of gis/ mapping experience so I cannot comment on the projections and handling of the shape file layers. Someone with more knowledge should give feedback on that aspect of the study.

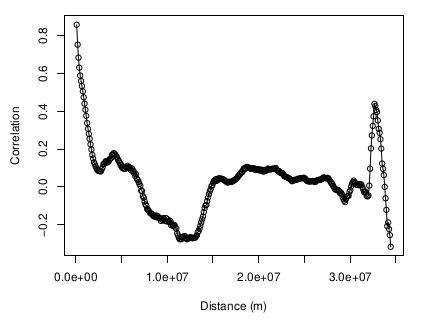
2) Bootstrap

While I am sure that the models are computationally intensive, I would like to see a single model run with a spatial autocorrelation term for each taxon so we can get test statistics and coefficient estimates. Larger models have been run and there are ways to deal with these issues (e.g., Ives et al. 2021).

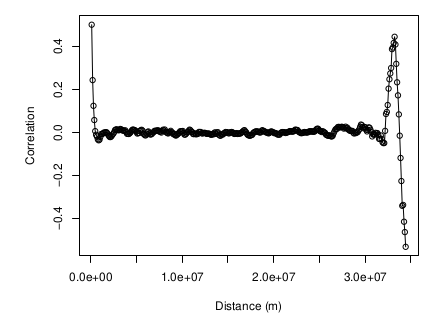
Ives, A. R., Zhu, L., Wang, F., Zhu, J., Morrow, C. J., & Radeloff, V. C. (2021). Statistical inference for trends in spatiotemporal data. Remote Sensing of Environment, 266, 112678.

**Thank you for the linked paper and accompanying package. Unfortunately, this was not possible for us to use as our data requires a binomial model. We do however agree that a more standard modelling approach could be used and were able to identify an appropriate method which was computationally feasible.**

**The description of the final method can be seen on lines 239 - 258, but in short, we used a Gaussian process smoothing approach utilising the bam function from the mgcv package which allowed for multithreading in the model itself. This significantly decreased spatial autocorrelation (which can be seen in the supplementary material in full, and we show an example in Figures A and B below). We also added random intercepts for ecoregion as this improved model fit and further reduced spatial autocorrelation.**

****

**Figure A. Correlogram of residuals for the full model for mammals without accounting for spatial autocorrelation.**

****

**Figure B. Correlogram of residuals for the full model for mammals accounting for spatial autocorrelation.**

3) Model construction and description

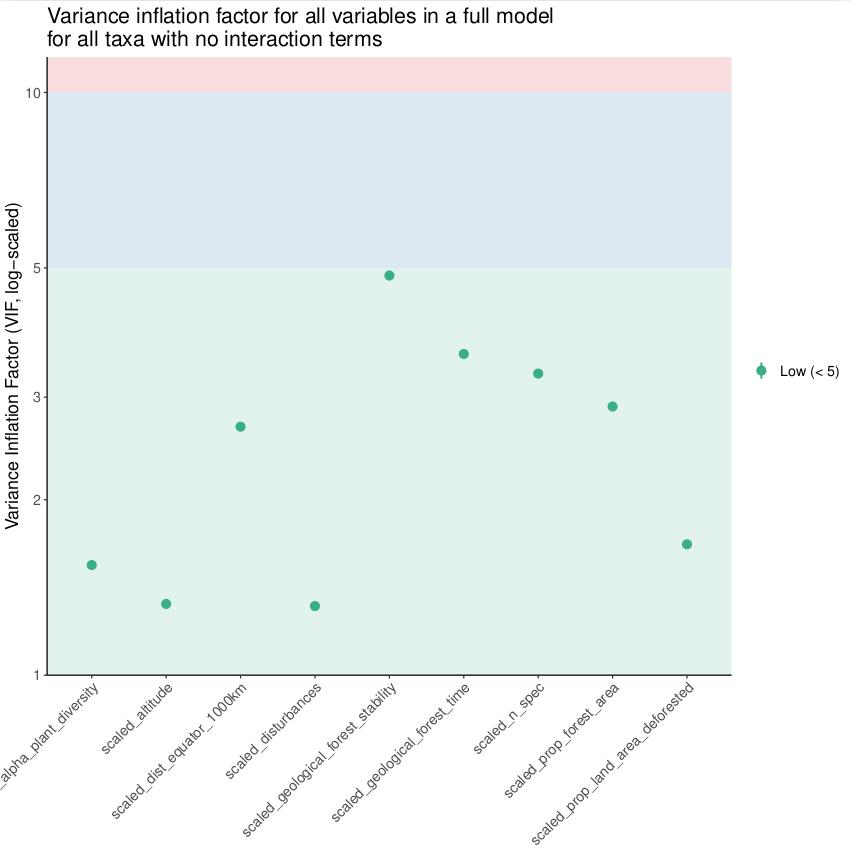
There is a lot missing regarding the model in the text. What distribution was used in the glm? What was the link function?

**We agree that more of the details of the model could be included in the full text. We now clearly state that we used a binomial model with the logit link function - lines 228 - 232 and write out the full model in the Model Structure section of the supplementary.**

**“We modelled the proportion of forest species for each taxon individually, as well as for all taxa combined, using two sets of explanatory variables. All models used binomial errors and the logit link function, since our response variable, the proportion of forest species is a set of trials (total number of species in a pixel), and successes (the number of forest specialist species in a pixel)”**

Definitely several of the variables were colinear, how did the authors deal with that?

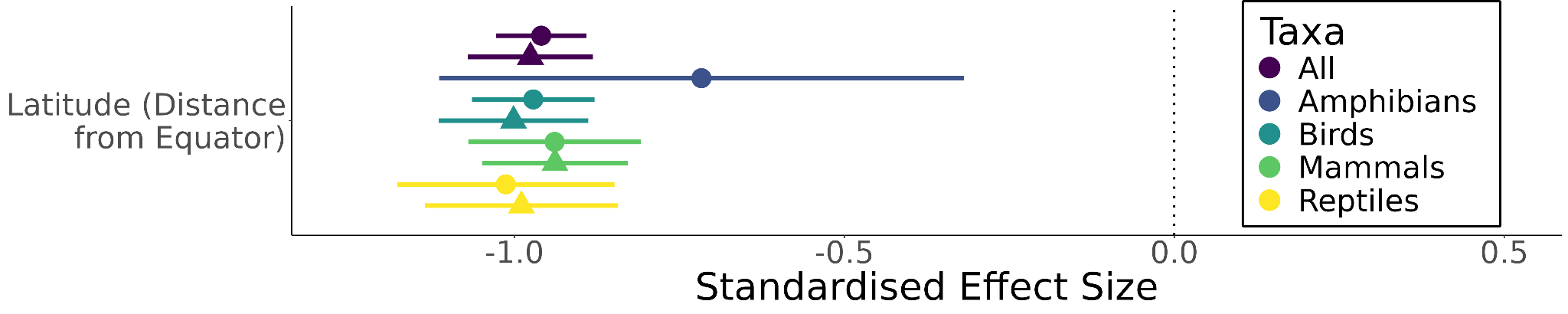
**We tested for collinearity using the performance package, as stated on lines 221 - 222. Using VIF scores we found that none of the variables were highly collinear. Since VIF plots per model were almost identical we report the VIF plot for a full model (including all explanatory variables) for a model including all taxa (and a random intercept for taxa). This can be seen in S1, and is also shown below in Figure C.**



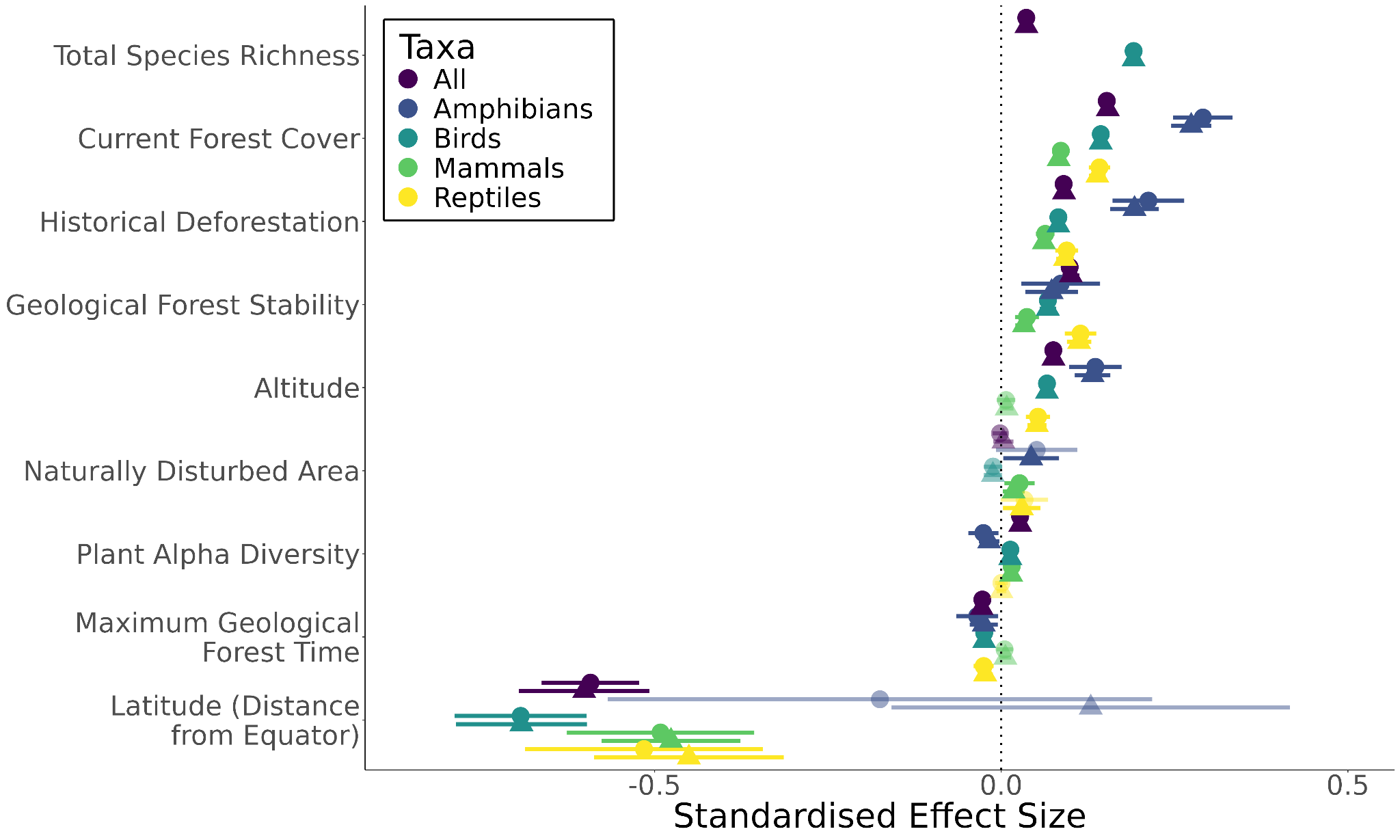
**Figure C. VIF scores for each factor in a full model without any interaction term.**

Were the model assumptions met using the tests in the DHARMa package (Hartig, 2022)?

**The model assumptions were broadly met and were tested using the testUniformity, testDispersion, and testQuantiles functions in DHARMa. All model diagnostics can be seen in S1. Since we had a large dataset (and therefore lots of residuals) we assessed model fit by visual observation. Models generally had slight overdispersion, though the degree to which they were overdispersed was very minimal and this was unlikely to affect results. To ensure this, we ran models using the quasibinomial family which accounts for overdispersion and found almost identical results for our models using the binomial family (See Figures D and E below). We kept the binomial results for the manuscript as other tests such as those in DHARMa and the performance package do not work for the quasibinomial family.**

****

**Figure D. Effect sizes of binomial (circle) models are very similar to those in quasibinomial (triangle) models, which account for the mild overdispersion seen in the binomial models. The amphibian quasibinomial model did not converge and is therefore not included.**

****

**Figure E. Effect sizes of binomial (circle) models are generally similar to those in quasibinomial (triangle) models, which account for the mild overdispersion seen in the binomial models. When large differences occur (effect of latitude on amphibians), effects are non-significant and therefore interpretation is the same.**

Why was the only interaction included the Current Forest Cover x Historical Deforestation? Justification is needed for this as I can imagine several interesting interactions.

**We agree that many possible interesting interaction terms could be included, between many of the variables included. We initially included an interaction term between the proportion of forested area and the proportion of deforested area as previous studies found that species in highly forested areas were more affected by deforestation than those in poorly forested areas. There is however likely equal justification for many other interaction terms.**

**As such we have decided to remove all interaction terms from the model. Whilst including other interaction terms could be of interest, they would retract from the tests of our main hypotheses; inferences of how a variable affects the proportion of forest species would be dependent on the value of the other interacting terms meaning interpretation is difficult, if not almost impossible when 4+ variables are interacting.**

How was the model fit?

**In terms of diagnostics, the model fit for all models included in the manuscript can be seen in the supplementary in S1. In terms of the variation explained, the R2 for all models is very high, however, this included the variation explained by the random effects and the smooth term accounting for spatial autocorrelation; since GAMs cannot be partitioned into conditional and marginal R2. Comparing the R2 between our models only including latitude and those including our other variables is however still useful, and these have been reported in Table 2.**

4) Elevation

Why wasn’t elevation used as a covariate in the context of Janzen’s Mountains are higher in the tropics? This goes back to the interdiction and the absence of certain hypotheses addressing specialization and, in the tropics, (and also evolutionary hypotheses are largely missing from the introduction)

**Thank you for the suggestion. We have added altitude to the model and have added this as a hypothesis in Table 1, as well as another interesting hypothesis that could link altitude (extreme environments) and the proportion of forest species. Additionally, we have tested for a further hypothesis which is not directly linked to the LDG but is very likely to be linked to the proportion of forest species; niche packing.**

5) Spatial autocorrelation

How does bootstrapping make up for not including a spatial autocorrelation term? I don’t understand how this would work and it is not explained.

**The bootstrapping was not used to account for spatial autocorrelation, but rather to allow us to run models using our initial method to account for spatial autocorrelation. Our new modelling approach does not require this and is explained on lines 246 - 258.**

6) Model output.

As I said earlier, I think the authors should run a single model and report model statics. There are no model tables (I realize you ran the mode thousands of times) or good ness of fit reported. At least report the average coefficients and test statistics across runs (could put in the supplement).

**Yes, we agree. Now we have run a single model per taxa we can report these in coefficient tables. These are reported in S2.**

7) Interaction plots

It doesn’t look like a significant interaction to me. The average lines for levels of deforestation should be plotted on in the same plot and more levels should be added. Further, a detailed explanation of the interaction is needed. What does “…amplifies their independent effects…” mean in the context of an interaction. What is the interaction coefficient and test statistic?

**We have no longer included this interaction term in our analysis to focus on testing our key hypotheses. As such the interaction plots have also been removed.**

Line 141-142: There has been a lot of deforestation since 2000 (the year of the forest map layer). While the biodiversity data may not be that sensitive (i.e., has been compiled over a longer period of time is not a fine spatial resolution) it would be good to at least mention this in the discussion, in my opinion.

**Thank you for the suggestion. We have added it to the methods section as we did not feel like it naturally fit in any of our discussion paragraphs. It is on lines 134 - 137, and reads: “While this data is now over 20 years old it still represents the best estimate of global tree cover, and is unlikely to affect results as the IUCN and BirdLife range maps were created over a similarly coarse temporal and spatial scale”**

Line 228: 10 % of what? Species or pixels?

**Pixels with less than 10% land area were removed, this has been rephrased on lines 224 - 226. We have also now changed the analysis to not remove pixels with a particularly low number of species, since the binomial models automatically weight observations by the total number of trials, which in our case is the total number of species. However, we have kept this for the plots since the smooth fit here does not weight by total species. This is explained in the legend for Figure 1.**

Lines 261-264: Unpack this “lagged” response a little more, it is interesting. Are there any citations for such an effect? Extinction debt?

**We further elaborate on this effect in lines 311 - 319 where we suggest, as you have done here, that it is possibly a case of extinction debt. We have further elaborated on this, adding information on how long this may take to be paid; centuries.**

Lines 290-293: I am confused about the logic of the “positive relationship between geological forest stability and the proportion of forest species” supporting the lack of relationship between proportion of forest species and the amount of time an area had continuously been the same forest. If I am interpreting this right, forest stability (forests remaining the same) amount of time an area had continuously been the same forest should be related and be in the same direction. The fact that they are not, is not support for positive relationship. I may just be confused…

**We agree that this could be explained more clearly. Maximum geological forest time, which has a very weak negative effect, is a measure of how long a pixel could have continuously contained the same forest type. We would expect this to have a positive effect on the proportion of forest species but it does not.**

**Geological forest stability is the amount of time periods (raster layers) in which the pixel contained the same forest type as it does now – the pixel did not have to continuously be covered by the same forest type. This was found to have a positive relationship with the proportion of forest species.**

**From the two above results it seems that the overall amount of time a pixel has contained the same forest type is important, but the amount of time it has continuously contained the same forest type is not. This suggests that forest species are either surviving in the pixel as forest types change, or, they are moving to refugia when the forest type changes and then moving back when it goes to the original forest type again.**

**We have attempted to explain this more clearly on lines 344 - 347, which reads: “This suggests that forest species are able to persist in areas when the type of forest present changes, either by moving to nearby refugia and repopulating when the forest type changes again to become favourable”**

Referee: 2

Comments to the Author

The paper “A global latitudinal gradient in the proportion of forest species” aims at identifying if the patterns in proportion of vertebrate forest specialist species follow the same latitudinal distribution than biodiversity gradients, and which factors might explain this pattern. For this purpose, the authors used range maps and habitat use data for terrestrial vertebrates to compute proportion of forest specialist species, and compare them to current forest distribution, historical deforestation, geological persistence of forests, disturbances and plant diversity, using GLM. The authors found that the proportion of forest specialist species is driven by current forest cover, historical deforestation, and forest complexity, and by a lesser extent by long-term habitat stability. Despite taking into account the effect of these variables, the effect of latitude was still highly significant, which suggests other underlying causes of this gradient that were not accounted for. Finally, the high proportion of forest specialist species found in deforested areas suggest that the response of these species is lagging.

This paper is well-written, clear and well organized; I enjoyed reading it. Among more minor specific comments listed below, I suggest the authors to be more straightforward with the comparison between LBG and there expected results. Right now, it seems a bit obvious that biodiversity gradients and proportions of specialized species will go in the same direction, that is why the authors should explain in more details why they study these proportions, why this is important, how it may differ from biodiversity distribution.

**We agree that the proportion of specialist species should generally increase as species richness increases (Figure 2 shows this relationship), however as expected this is not a perfect relationship and there is considerable deviation to be explained. We have emphasised how our results differ from the LBG in species richness by including species richness as a predictor in our models, and including information on why we may expect areas with higher species richness to have more specialist species in Table 1, and explained this further in the discussion (lines 320 - 328)**

**Additionally, we believe we have now better outlined why studying proportions could be important for conservation (above and beyond just using species richness or other diversity metrics), in the first paragraph (lines 42 - 58).**

Biodiversity might not be enough, and studying specialist species brings more information on how these species might be impacted by habitat change and degradation.

**Thank you for this comment. We have almost completely rewritten the first two paragraphs of the introduction to emphasise why we should look at whether the LBG holds for species when thinking about their sensitivity to habitat loss - the proportion of specialist species. This is on lines 42 - 75.**

Also, this may have conservation implication, and a short paragraph in the discussion might be added to discuss this issue.

**Thank you, conservation implications have been added to the final paragraph on lines 395 - 402**

In the introduction and discussion, some paragraphs sometimes lack transition between them, and working on the flow and transitions might help the reader.

**We have re-written the first two paragraphs of the introduction which should help transition between future paragraphs by giving more of an overview of what we are doing and why we are doing it.**

Otherwise, the figures are nice and clear, with necessary information. Maybe adding an x axis on figure 1 A-D2 would help the reader.

**Thank you, we do not believe that Figure 1 would benefit from an x-axis since we do not include longitude in our analysis, and therefore the specific value of the x-axis is not important for the reader. Figure 1 is primarily for readers to see the clear latitudinal gradient present.**

Specific comments.

Title: I would suggest to be more specific, as forest species often refers to plant species, and add “terrestrial vertebrates forest specialist species”; or something similar.

**Thank you for the suggestion, we have amended the title to “A global latitudinal gradient in the proportion of terrestrial vertebrate forest species”**

Abstract. I would suggest to straightforwardly explain why studying the proportion of forest specialist species is important, what does it bring compared to biodiversity indices, and why we could expect a different pattern from the ones observed for biodiversity gradients.

**We have added a sentence about why studying the proportions of species is important, and how it could uncover different results/implications to the standard LBG. This can be seen in lines 9 - 15.**

Introduction.

L.46-48: Maybe emphasize that this is about specialist species (and why are they different from generalist ones), and also develop why this is important and you expect a different pattern than the one observed for biodiversity. Better articulating the two first paragraphs might help understand what is at stake here.

**We agree that adding more information about why it is important and interesting to look at how the proportion of forest specialist species varies globally would be beneficial. As such we have rewritten the first two paragraphs, to highlight this from lines 42 - 75.**

L. 65-66: By “greater amounts of forest at regional scales”, you mean less fragmented, higher proportion in the landscape?

**This section has been removed as it is now shown in Table 1. In Table 1 current forest cover and historical deforestation are described in the proportion of the area as opposed to simply amount.**

L. 76: the “current” global distribution of forests alone.

**This section has been removed and moved to Table 1, and therefore no longer needs to be amended.**

The different factors potentially impacting the proportion of specialized forest species are well described in the introduction and provide a nice overview. However, I think the introduction would benefit from a paragraph that would put into perspective the expected combination or hierarchy of these factors, explaining that this is what you are going to test, before directly jumping into the datasets used (L.108).

**We have added an additional section in the final paragraph to summarise the hypotheses before then talking about the data and methods used to test the hypotheses. Lines 91 - 99, read: “In summary, we predict that the proportion of forest species will increase with the amount of current and historical forest cover in an area, as well as the stability and complexity of the forests, where we expect higher proportions of forest species in areas which have had few disturbances and have remained the same forest type. We additionally expect factors which increase specialisation and speciation to increase the proportion of forest species, specifically total species richness and altitude. A combination of these factors is predicted to create a latitudinal gradient in the proportion of forest species, which is likely to be most heavily driven by the disproportionately high amounts of forest at the equator.”**

Methods.

L.125-126: is there a link to access these datasets?

**The final dataset used for analysis is available on Figshare (**[**https://doi.org/10.6084/m9.figshare.23798829.v2**](https://doi.org/10.6084/m9.figshare.23798829.v2)**) however we have not uploaded the IUCN and BirdLife ranges since these are already freely available for non-commercial use. A text file with links to the IUCN and BirdLife webpage where range maps has been added to Figshare (**[**https://doi.org/10.6084/m9.figshare.25053560.v1**](https://doi.org/10.6084/m9.figshare.25053560.v1)**).**

L. 131-132: was there an evaluation of how well these datasets are categorizing specialist vs generalist species in the literature?

**We do not know of an independent global dataset which could be used to test our categorisation. Datasets for individual taxa exist, for instance, AVONET, however, these classifications of habitat use data from IUCN/BirdLife which is used in this study, and would therefore not be independent of our dataset.**

L. 143: Is there a reference for justifying the use of a cut-off value of 70%?

**70% tree cover was chosen as a conservative categorisation of a forest, as classification of what constitutes a forest varies nationally, yet for our analysis, we had to ensure forests were consistently categorised globally (e.g 60% of tree cover is determined as forest in Brazil’s mapBiomas, 20% in the United States USGS National Land Cover Database, and 70% in the Copernicus Global Land Cover Dataset).**

**This has been explained in text on lines 139 - 143.**

Also, with this method the size of the forest fragments was not evaluated. Is this important?

**We agree that patch size and other fragmentation metrics such as connectivity could be of interest. We would argue that using a 70% cut-off for pixels incorporates patch size and connectivity better than using another method such as averaging tree cover. By using a cut-off of 70% tree cover for the 30m pixels we are ensuring that the pixel is highly forested. This means that when we calculate the proportion of 30m pixels in our larger 96km pixels we can be sure that this is the proportion of highly forested area. The alternative would be to calculate the mean tree cover in the 96km pixel by taking the average tree cover of the 30m pixels, this could be quite high even if the region has no forest; for example an area with savannah could have 20% average tree cover, but using our method it would have 0% forest cover since none of the individual 30m pixels would be calculated as forest (none have >70% tree cover).**

**Using our method and at the scale of our analyses, the proportion of forested land area should be a relatively strong proxy for patch size and connectivity; higher proportions of forested land mean larger patches and higher connectivity.**

L. 153: even if the dataset is from another paper, more details are needed to understand how this map was generated. Please add information.

**We agree, and we have added additional information to lines 153 - 156, which read “This map was generated using a random forest regression model which predicted the distribution of historical forest as a function of 19 bioclimatic variables from the WorldClim database, and the coordinates of forest loss from the Hansen tree cover dataset”.**

Also, this is not clear how the deforested area was computed here, it is said “the total deforested area was calculated by summing the area of the forested pixels”, but you said you used a map of forest cover prior to human activities. First, did you subtract historical forest cover to the current forest cover?

**We agree this could be better explained. The map we used had already removed the areas of current forest cover and thus was a map of areas which were predicted to have been forest but are not currently forest; deforestation. This has now been explained in lines 150 - 153 which read:**

**“The historical proportion of deforested land area was calculated using a map of predicted forest cover prior to human activities, which had had the areas of current forest cover removed, this map was also already categorised as forest or non-forest”**

L. 157-163: please clarify this part.

**The initial pixel size of the historical deforestation raster was different (coarser) to that of the forest and land raster in the Hansen dataset. As such, on rare occasions, when we calculated the proportion of deforested area (deforested area/land area) the value was greater than 1. We set all pixels which had a proportion of deforested area greater than 1 to a pixel value of 1, since more than 100% of an area cannot be deforested. This is described in lines 150 - 167. This reads:**

**“The historical proportion of deforested land area was calculated by dividing the total deforested area by the total land area. Note that due to the different initial pixel sizes of the land area raster and the deforested raster some pixels contained values greater than one. Most of these were removed during the ‘Data Preparation’ process described below, but 85 of 5,434 pixels for amphibians (1.6%, mean value > 1 (sd) = 1.13 ±0.317), 125 out of 9,571 pixels for birds (1.3%, mean value > 1 (sd) = 1.13 ±0.305), 125 out of 9,435 pixels for mammals (1.3%, mean value > 1 (sd) = 1.13 ±0.305), and 101 out of 6,125 pixels for reptiles (1.6%, mean value > 1 (sd) = 1.05 ±0.087) remained with values above one, and as such these were all set to a value of one, since the proportion of land area deforested cannot by greater than the total land area present”**

L. 165: “historical naturally disturbance”, did you mean historical naturally-occurring disturbance?

**Yes thank you, this has been changed.**

Here again, more information is needed on the origin of the dataset used. Also, it would be interesting to know what are the expected impacts of each of these disturbance on forest species, as they are expected to differ.

**We have added more information on how this data were generated in lines 169 - 173. In our analysis we do not assume they will differ as we use the number of disturbances as our explanatory variable, therefore assuming fires, storms, and glaciation have the same impact on the proportion of forest species, and that what matters is the number of these disturbances. This now reads:**

**“Maps of historical naturally-occurring disturbance areas caused by fire (Lavorel et al., 2007), glaciation (University of Geneva, Switzerland et al., 2001), and storms (Location of tropical cyclones) were extracted from relevant sources where they were generated by expert opinion (University of Geneva, Switzerland et al., 2001; Lavorel et al., 2007) and satellite observation (Location of tropical cyclones)”**

L. 175: maximum geological forest time?

**This makes more sense and the variable name has been changed accordingly**

L. 201: you can’t say that comparing plant diversity between forest and grassland has no ecological meaning (in general!), but in the context of your study this is true. Please rephrase accordingly.

**We agree and have made this clearer on lines 210 - 214. This now reads:**

**“For example, differences in plant diversity between two forest ecoregions suggests changes in structural complexity and the number of niches available to forest species. On the other hand, comparing differences in plant, and therefore structural, diversity between a forest and grassland ecoregion has far less ecological meaning in the context of the proportion of forest species.”**

L.212: “two sets of explanatory variables”.

**Thank you, this has been changed.**

Please explain how the significant explanatory variables were selected in the model.

**We chose explanatory variables based on our hypotheses and did not do any model comparison based on AIC to remove variables and find the “best model”. We have now added a sentence explaining this to lines 242 - 245, which reads “The structure of all models can be seen in the ‘model structure‘ section of the supplementary, and model variables were chosen based on our prior hypotheses and were not included or excluded based on model comparisons such as AIC”**

Also explain in this section ho the “standardized effect size” presented in the figures is computed.

**The standardised effect size is the base output of the model since our variables have been scaled and centered, which is explained in lines 238 - 239 which reads “In both cases, we ran generalised additive linear mixed models, with explanatory variables scaled and centred to allow for comparison of effect sizes”. This has now been made clear in the figure legends of Figures 3 and 4 which now include “All variables have been scaled and centered to standardise effect sizes and allow for comparison between variables and models.”.**

Results

Maybe no section title is necessary here as there is only one.

**That makes sense, it has been removed.**

L. 246: please add a bit more here about Figure 4.

**We have removed the interacting effect of current forest cover and deforestation and therefore figure 4 has also been removed. This was done to focus the manuscript on our main hypotheses.**

Discussion

L.268-269. Maybe working on the transition can improve the flow of ideas. Before you summarized your results, and then you discuss in more details the different factors. “The first obvious factor we tested was the amount of forest cover and as expected…”.

**We believe that we have already used this approach of summarising our results and highlighting the most interesting findings in the first paragraph (lines 290 - 304), before then going into more detail about each hypothesis in either individual paragraphs or combining related hypotheses in a single paragraph. This approach is suggested in scientific writing guides (e.g. Writing Science by Joshua Schimel) and as such we think it is appropriate.**

L. 281: is there some papers that studied how long we may wait until the debt is paid? Also, this has conservation implication. Should describe those.

**The conservation implications have been described further on lines 400 - 402, and reads: “Additionally, our results suggest that reforesting historically deforested areas is a conservation priority since they contain disproportionately high proportions of forest species, which may be declining towards local extinction”**

**We had similarly added a sentence outlining how long such extinction debts could take to be paid (lines 316 - 319), which reads: “which only increase as the size of the area increases (Haddad et al., 2015), meaning that this may be a case of extinction debts that are yet to be paid and could take centuries to do so (Chen & Peng, 2017)”**

L 293-296: references to support this?

**Added references about refugia during the last glacial maxima which is probably the best example of this.**

L.314: Please explain in SI how the hand wing index presented was computed.

**This has now been explained in the SI as follows “Hand wing index is a well-studied proxy for dispersal ability (Sheard et al., 2020; Tobias et al., 2022; Weeks et al., 2023), and is calculated as the ratio of the Kipp's distance (the distance between the tip of the first secondary feather and the tip of the longest primary feather) to the total wing chord (Sheard et al., 2019)”**

L. 318”a relatively recent analysis”, avoid relatively.

**This has been removed**

L. 319: how is your classification more specific?

**It is more specific to our question since it groups species into forest or non-forest as opposed to open or closed area. This has been made clear on lines 370 - 373 by simply removing more specific and reads:**

**“Whilst a recent analysis found no significant differences in the hand wing index of closed versus open-area species (Sheard et al., 2019), analysis of our categorisation of species into forest or non-forest suggests that forest species are indeed poorer dispersers”**

L. 323-325: not clear, please elaborate.

**This has been removed as it was not clear and results from our new analysis do not support it.**

L. 345: reference?

**Added references.**

Here it would be nice to have a paragraph that synthesis how your results differ from the LBG hypothesis, why using forest specialist proportion is important compared to biodiversity data only, and the potential impacts of you results for conservation implication.

**We agree, this has been combined with our current final paragraph on lines 395 - 402.**