**Annotations concerning the review of “Challenges in estimating species ages from phylogenetic trees”**

***Main common issues:***

* The mathematical rigor of the equations proposed for the age estimator.

*Following reviewer#4 manual to improve and describe our equations and/or implementing the reviewer#3 suggested method.*

* The species concept philosophical dilemma.

*We have a text about the species concept on previous versions of our manuscript. I think we can use them, probably in the introduction. I suggest that we do not enter deeply in this subject, because is a philosophical mine field.*

* The biogeographical relevance of our paper. Linked to the editor’s concerns regarding the adequacy of our paper to GEB.

*I think, regarding the suitability concern, we should argue why we chose GEB. I recall that we based our decision on recent GEB publications that were using the phylogenetic age approach. Also, our paper is a critic and a partial solution for macroecological research linking age with other variables, and mostly macroecologists conform the GEB readership.*

* Reproducibility.

*We can solve this problem by setting our GitHub repository as open access.*

***Major Critics:***

**Editor**

* “I therefore encourage the authors to consider very carefully reviewer comments on the attached pdf along with the more technical comments from reviewer #3, and to provide a more detailed and transparent description of their algorithmic approach.”
* “More thorough consideration of species concepts”
* “I would encourage the authors to think of alternative methods to evaluate their age estimates. Perhaps there are some empirical phylogenies that include extinct taxa”

**Reviewer #1**

* Species Concepts: The manuscript would benefit from a more detailed introduction to the species concept that implicitly underlies the "true species age" you refer to. The idea of a “species age” itself relies on considering morphological/vertical species (with associated arbitrariness) and can seem confusing to some biologists accustomed to the biological/horizontal species concept.
* Validating the Probabilistic Method: I would suggest a further validation analysis to convince the readers of the method’s reliability. This could be a full calibration curve based on the posterior distributions of estimated species ages or simply the 95% coverage.
* Reproducibility: It is crucial that all necessary materials for reproducing your results are available. The 'source.R' file was not found in the GitHub repo, and I was therefore unable to replicate your results. I would also suggest providing either the CSV files or the order in which to run the scripts to recreate them.

**Reviewer#2**

* Species concept: A fourth shortfall could be related to species concepts and the philosophy of biology. What is a species in the context of time? When could a species be said to have originated? Do species even have ages, or are they just a contemporary phenomenon and before the present they are simply ‘lineages’ because we lack the information to define them? This is related to shortfall three about speciation mode and I’m not sure whether it represents its own shortfall, but could be worth mentioning. An argument could be made to rethink the idea of species ages altogether (or completely drop it), and I think your research here puts forward a reasonable case for that.
* Biogeographical relevance: One context related comment is that it is not exactly a biogeography paper. While I don’t think this is hugely problematic because it is of extremely high relevance and value to biogeographers, perhaps the discussion could highlight some areas of biogeographic research that may be influenced by this research. You talk about extinction risk but the most obvious that comes to mind is the estimate of assemblage ages based on the mean or max phylogenetic species age metric. Lots of studies do this (I’m guilty of it as well) and there are many reasons why it doesn’t make much sense.
* Species age-Extinction risk: This part is not really discussed in any detail in the discussion section. Although not monotonically increasing in all cases, it still does increase positively. Therefore, if you measured age on a continuous scale (as is done in most studies) you would see a generally positive relationship. Although the slope of this line will be less than if you had the true estimates, you will still reach a similar conclusion about the effect. So, I wonder whether this conclusion is overplayed a little? Is it worth discussing the extinction risk analysis in more detail? This part could be linked to my previous point about consequences for biogeographic studies.

**Reviewer#3**

* Mathematical rigor: Concerns about the clarity and the description of each equation. However, the reviewer also gives their own method to estimate species ages. We should test it. Think about "better than the worst method" does not mean "a good method".
* Separate our age estimator from the simulation study. It would give more space to go into the mathematical details.

**Reviewer #4**

* Mathematical rigor of the equations proposed. However, the reviewer gives a step-by-step manual of how to improve our equations and their descriptions. I found it very gentle.

***Minor comments:***

**Editor**

To ensure that your paper conforms with GEB style guidelines, please check the following:

1. A biosketch and a data availability statement should be included.

2. A short running header should be included to go at the top of the journal page.

3. Citations in the text should be in house style.

4. English spellings should be used throughout the text.

5. The reference list should be in house style.

6. All the citations in the text should be listed in the References.

7. The figures and tables should be in numerical order.

8. A short title (not a complete legend) should be supplied for each numbered item in the supplementary material. Please place these at the end of the article after the figures and just before the supplemental material itself.

**Reviewer#1**

1. \*\*Suggestion to Replace Fig. 6 with Fig. SM5\*\*: It seems easier and more informative to me to directly compare the MAPEs side by side instead of using ΔMAPE.

2. \*\*Clarify Visualizations\*\*: I suggest considering adding barplots to Fig. 3 & SM1 to show proportions of under-/well-/over-estimated species age, otherwise the high proportion of congruent points (e.g., for bifurcating speciation at low extinction fractions) is not visible.

3. \*\*Clarifying the Second Approximation\*\* (l.191): The rationale for this second approximation might be intuited, but please articulate it in the text.

4. \*\*Typos\*\*: Please correct the double comma (l.113), double dot (l.171), and missing space between “tau” and “did” (l.212). On l.291, “exceeding 25%, irrespective of the extinction level” is not strictly true, as the error rate is 24.4% in the absence of extinction.

**Reviewer#2**

* Line 101-104 – this is a good example of where the definition of species age is really blurry and highlights how it is a problematic concept. Especially, because diversification models do not include rates of anagenetic vs cladogenetic speciation.
* Line 113: there are two commas after ‘sampling’
* Line 117-127: Is it worth giving a brief description of TreeSim and FossilSim algorithms for context so the reader doesn’t have to completely refer to other papers.
* Line 140: indent this para?
* Line 160: budding and bifuricating in equal proportion?
* Line 166: how many randomly selected pairs do you compare from each tree?
* Line 171: two . After 0.25
* Line 195-196: This sentence doesn’t quite make sense.
* Line 199: So you need to know the sampling fraction beforehand? What if you don’t have this info for cryptic or undescribed species?
* Line 212: space after t
* Line 246. I don’t see how Figure 3 shows this point as both figures contain some extinction right?
* Line 257-258: This is inconsistent with the statement above “At low extinction fractions (< 0.25), 96% of the phylogenetic age estimations were congruent with the true age”. What does congruent mean in this context?
* Line 269-270: Is this because all trees simulate a fixed number of species? What about the net diversification rate as a useful metric here because of the scaling of the tree ages.
* Fig 1. Is not essential to the manuscript so if you need to create space I’d recommend that. Ignore if you don’t.
* Fig 3. Would plotting the line of best fit for each extinction scenario be helpful for visualizing the different trends here?
* Fig 4. Same point. Would trend lines be helpful here. Relationships look positive but quite hard to tell. Also note that the figure is not green-red colour blind friendly.
* Line 384: and its possible budding speciation (e.g., dispersal based) might be the most common mode in vertebrates (Skeels and Cardillo 2019, already cited)
* Line 387-390: Can you explain why it is at odds with birth death models? Seems like an important point.

**Reviewer#3**

None.

**Reviewer#4**

• Line 98-99: it is also necessary to have 100% taxon sampling for this to be true

• Line 113: double comma

• Line 171: double period