**Review and responses for “Challenges in estimating species age from phylogenetic trees”.**

Dear Editor,

We highly appreciate your time and dedication going through our manuscript, which have now thoroughly revised in the light of your comments.

Explain main modifications to the manuscript

Please find below our point-by-point replies addressing each issue raised by the editor and the four reviewers.

Editor Jonathan Davies

1. There was broad support among reviewers for this manuscript, although some significant concerns were raised, and it was also not clear to me whether GEB was necessarily the best venue you for it, as also hinted at by one of the reviewers, but I defer to the editorial board on this.
2. We managed to obtain four in-depth reviews, and so I will not burden the authors with additional critiques; however, I note that the manuscript addresses an issue that anyone using phylogenetic trees to infer species ages should be intimately aware (or, arguably, they should not be analysing phylogenetic data), and so the key advance here is in providing a meaningful solution to the challenge of estimating true species’ ages from phylogenetic trees of extant species - the birth-death-based age estimator. 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.
3. Key points that would need to be addressed in any revision additionally include a more thorough consideration of species concepts, and 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. As noted in the reviews, simply showing the method performs better than a method we know is bad is a low bar.

Referee 1:

1. Your manuscript "Challenges in estimating species age from phylogenetic trees" highlights the important issue that attempting to infer species age from phylogenies is seriously undermined by incomplete sampling, hidden extinctions, and non-bifurcating speciation modes. The proposed probabilistic method for estimating the age of species is interesting and useful, even though it is applicable only under the assumption of bifurcating speciation. I believe this article should be accepted for publication, conditioned on providing the code currently lacking for replicating the results.
2. \*\*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.
3. \*\*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.
4. \*\*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.
5. \*\*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.
6. \*\*Clarifying the Second Approximation\*\* (l.191): The rationale for this second approximation might be intuited, but please articulate it in the text.
7. \*\*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.
8. \*\*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.
9. I thank you for this work and look forward to the revised manuscript with these points addressed.

Referee 2:

1. I have read and reviewed the manuscript “Challenges in estimating species age from phylogenetic Trees”. I was very excited to see a paper on this topic and the authors put together a very convincing argument for why estimates of species ages may be flawed using current methods while also presenting a new tool to better estimate them. I found the paper very well written and I don’t have any major criticisms that should prevent the publication of this manuscript (although I am not confident in my ability to thoroughly review the details of the analytical model, what I understood made sense). Though I do have a few comments. These should be taken as recommendations but I’ll leave it up to the authors/editors whether to address or not.
2. Introduction - 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.
3. 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.
4. The relationship between age and conservation status. 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.
5. Minor comments
   1. 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.
   2. Line 113: there are two commas after ‘sampling’
   3. 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.
   4. Line 140: indent this para?
   5. Line 160: budding and bifuricating in equal proportion?
   6. Line 166: how many randomly selected pairs do you compare from each tree?
   7. Line 171: two . After 0.25
   8. Line 195-196: This sentence doesn’t quite make sense.
   9. Line 199: So you need to know the sampling fraction beforehand? What if you don’t have this info for cryptic or undescribed species?
   10. Line 212: space after t
   11. Line 246. I don’t see how Figure 3 shows this point as both figures contain some extinction right?
   12. 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?
   13. 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.
   14. Fig 1. Is not essential to the manuscript so if you need to create space I’d recommend that. Ignore if you don’t.
   15. Fig 3. Would plotting the line of best fit for each extinction scenario be helpful for visualising the different trends here?
   16. 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.
   17. 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)
   18. Line 387-390: Can you explain why it is at odds with birth death models? Seems like an important point.

Referee 3

1. Understanding how old extant species are is critical to addressing both basic and applied questions in evolution and ecology. The authors rightly point out that a common method for estimating species age from phylogenies---simply taking the age of the node ancestral to the tip---is problematic. This estimator, which the authors call the "phylogenetic age", is either an overestimate or an underestimate of the age, depending on the mode of speciation, and on rates of extinction and sampling. The authors propose a new method for estimating species ages based on birth-death processes; they perform a simulation study showing how the phylogenetic age estimator is biased, and demonstrating that their new estimator provides more reliable estimates of species age.
2. The manuscript is very well written, establishing the problem clearly and arguing/demonstrating persuasively that there is a better way. I read this manuscript with great interest, as I have personally encountered people using the phylogenetic age estimator and myself made similar (though less thought out) criticisms in personal communication and reviews. However, this point has not been made clearly in the literature, and it will be nice to have a citation to point to, as well as a concrete recommendation for a better approach! Nonetheless, I have some technical concerns about the new method itself, and a minor comment regarding the structure of the manuscript. I detail these concerns below. Regardless of the criticisms, I find this work to be very thoughtful and interesting, and I hope that the authors take my comments in the constructive spirit in which they are intended! I am very optimistic that they will be able to address my concerns adequately for publication in Global Ecology and Biogeography.
3. The mathematics of species age under a birth-death process
   1. Starting on line 178, the authors provide a method for estimating the age of an extant species in a phylogeny based on a birth-death processes with bifurcating speciation. The exposition of this method is a bit unclear, and I am concerned that there are some technical problems.
   2. The authors begin with the equation for one descendant given one ancestor over an interval of length t (line 188), but use it as an approximation for the probability of no speciation or extinction over an arbitrarily small time bin of size t. Of course, for very small t, this is a good approximation of the probability of no events! But so is 1 - (lambda + mu) \* t. Indeed, when the authors started talking about arbitrarily small time bins, I was immediately expecting them to use difference/differential equations to describe the probability of no events over an interval, which would naturally begin with the equation I just provided. However, it seems that they are just using binned calculations in a numerical procedure, which was not at all clear from the start.
   3. The next equation on line 192 is confusing because (again) the numerical procedure is implicit, but my understanding (based on the supplemental code) is that the authors computationally divide the interval between v (the age of the ancestor) and the present into small bins of duration t (such that there are v/t bins), then compute p1 (the probability of no events) for each bin, then take the product over intervals as the probability of no events over the entire duration. Certainly the intervals are independent, so this seems like a fairly rational thing to do, but it all seems quite ad hoc/approximate. Normally, we present equations we want to solve, and then either solve them analytically (if possible) or numerically (if necessary), rather than jumping straight into an approximation.
   4. The next equation (equation 1) on line 198 is hard for me to follow. Where does the first term/ratio come from? Why do we compute the probability of survival of the sister, when it is a given that the sister survived (we are not computing the probability of the subtree, we're computing the probability that the node produced the extant species, given the subtree)? Overall, it is not clear to me what the logic/justification for this equation is. Similarly, equation 3 (line 207) is presented without derivation or much explanation. Again I have the feeling that the approach here is ad hoc, that the authors are intuiting their way through some math and/or developing a numerical procedure that seems reasonable, but is not well explained or strongly justified. At the very least, the probability that the true age of a species is at time t should include the rate of speciation at time t, lambda(t). Of course, the authors are explicitly assuming a time-homogeneous model, so perhaps this term would cancel out somewhere, but I can't tell because we're not given any algebra.
   5. Because I had difficulty following the presented equations, I attempted to derive my own. The probability density that the species age is exactly t should be: the probability that there is a speciation event at time t, times the probability that one daughter went extinct/was unsampled, times the probability that the surviving lineage experienced no hidden speciation events (events that produced only one daughter with extant descendants). I work through this math in an attached pdf, with code/simulations demonstrating that my equations work. I don't know if my results are comparable to the authors', because I did not try to recreate their numerical procedure, but notably I can compute the exact probability density that the age is t without using any numerical approximations. (Computing the mean/median of this distribution seems more difficult, so I did that numerically too.)
   6. To demonstrate the performance of their estimator, they compare it against the behavior of the phylogenetic age estimator. While I think that comparison is critical to the overall point of the manuscript, it doesn't actually demonstrate that the new estimator is correct or approximately correct. Given that the phylogenetic estimator is basically a worst-case-scenario for a bifurcating speciation model, pretty much any estimator should get you closer to the truth! For example, maybe the estimator v/2 performs better than v. My point is that "better than the worst method" does not mean "a good method". The authors should do a simulation targeted at demonstrating that the equations they present do actually recapitulate the distribution of species ages (or else use different equations that do get the right distribution). I've provided at least an example of such a simulation in the attached code. Of course, the work that I've done could be incorrect. But at least the authors need to do a better job of describing where their equations come from, and how their method works computationally.
4. The new method is hidden inside the simulation design. This is a more minor comment, but I find it a bit disappointing that the main novelty of this manuscript---the birth-death-based age estimator---comes so late, and in the middle of a section that is otherwise describing a simulation study. I think that the new method deserves more space, and as I mention above, more exposition. My recommendation is that the authors separate the new theory/method into a separate section before going into details of the simulation study. This will give them more space to go into more detail about how their method works, maybe even with figures and algorithms, as necessary.

Referee 4

1. Let me begin by saying that I very much enjoyed the narrative of the manuscript, and the presentation of the issue with treating divergence times as ”species ages”. I was not aware of this practice in the literature, and as the authors point out, this practice seems dubious. If anything, the critique of using phylogenetic divergence times in this way could have even been made stronger. The narrative is clearly written, and the discussion beautifully lays out the drawbacks and advantages of the proposed solution to estimate species ages from phylogenetic trees. Especially lines 430-433 explains nicely that phylogenetic age is only a good proxy for species age under specific, unrealistic circumstances. I completely agree with the authors that some correction is in order, in order to avoid biased estimates of species ages.
2. The authors propose to correct for this potential bias with a novel probabilistic approach. With this specific approach, however, I’m less convinced. I found that the probabilistic approach was either not based on first principles, or was not explained sufficiently. Since unsampled or extinct lineages are not included in reconstructed phylogenies, it is intuitive that divergence times should be older than species origination times. However, it is not clear exactly what kind of correction should be used, or what kind of model is to be used. Even if an incorrect model is used, or no model is used at all, it seems intuitive that any kind of reasonable correction will help reduce estimation errors in species ages. The authors discuss several modes of speciation, which is interesting, but my concerns are mostly about the derivations of the probability (densities) of the species ages. This, in particular, seems not sufficiently motivated or explained, and must be revised before publication can take place. Because I’m writing mostly about the equations, I decided to write this in LaTeX. I hope it is alright.
3. With this in mind (the detailed mathematical explanation provided by the reviewer) either give titles/names to the parts of equation (3), for example as I did above with overbraces, or break the equation into several smaller equations with appropriate descriptions.
4. explain why is the ”extinction probability” (1−P) being used here, whereas before it was the ”survival probability” (just P)?
5. fix their probability density function such that it returns positive numbers
6. ensure that it integrates to one across the data (i.e. τ ∈ {0, v}), or give a convincing argument for why this is not necessary (I may have misunderstood, and τ should be allowed to be older than the divergence time, in case of budding speciation mode?)
7. Minor comments
   1. Line 98-99: it is also necessary to have 100% taxon sampling for this to be true
   2. Line 113: double comma
   3. Line 171: double period