

Dear *Biogeosciences* publication team,

Thank you for the opportunity respond to reviewer comments posted about our manuscript entitled “**Quantifying the influence of wood carbon fractions on tree- and forest ecosystem-scale carbon estimation in a temperate forest.**” We appreciate the reviewer comments, which we found both supportive and constructive. Notably, all comments have been addressed in our revised manuscript.

***Biogeosciences* Reviewer: 1**

Overall, this is an outstanding work with a clearly defined research question, rigorous experimental design, in-depth and thorough analysis, and conclusions with significant scientific and applied value. The paper clearly reveals the superiority and necessity of using species-specific wood carbon fraction data relative to various common assumptions (such as the default 50% or the IPCC recommended value) when estimating forest carbon storage based on biomass. This fills a gap in the quantitative assessment of errors from single tree to ecosystem scale, providing crucial empirical evidence for improving national and global forest carbon monitoring and modelling predictions. The paper is well-structured, with detailed data, clear figures and tables, and strong arguments.

- Thank you for this kind endorsement of our research.

Although the overall quality of the paper is high, there are still some aspects that can be considered for further refinement in subsequent revisions or future research.

1. Caution in extrapolating the research site. The paper is based on a single temperate forest plot in Ontario, Canada. The authors have pointed out in the discussion that the carbon fraction of the dominant gymnosperms (such as balsam fir) in this plot is close to 50%, which may make the error of the 50% assumption relatively "mild" in this study. When extrapolating specific error margins to temperate forests with vastly different species compositions or even at a global scale, this context-dependent nature needs to be further emphasized in the discussion, and future validation in different types of forest ecosystems should be recommended.

- Agreed. In a revised manuscript, we wish to retain some extrapolation of our findings, though will explicitly note that species composition (among other factors) are reasons to be cautious in such exercises. Additionally, we note that including such content will indeed motivate further research that expands our studies to multiple forest sites that differ in species composition, stand structure, and other factors (Page 10, Lines 393-397).

2. Completeness of Carbon Fraction Variation. The study primarily focused on interspecific variation and reasonably cited literature to explain that this is the main source of variation. Intraspecific variation (such as tree age, location, and environmental gradient) was well reviewed in the discussion section, pointing out that its relative contribution is small but not yet fully clarified. This can serve as a clear direction for developing of forest carbon cycling models.

- Agreed. In our revised manuscript, we have more explicitly noted that variation that exists within species across different woody tissue types such as bark and stem wood tissue

differences, would likely be the best path forward in terms of accounting for intraspecific variation into carbon accounting models and protocols (Page 12, Lines 455-463).

3. Comprehensiveness of the Carbon Pool. This study focuses on the aboveground living biomass carbon pool (accounting for ~38% of the temperate forest carbon pool). It can be briefly mentioned in the discussion section that the wood carbon fraction assumption may also affect the estimation of other important carbon pools such as deadwood and underground biomass, which is also an aspect that needs to be considered in comprehensive wood carbon fractions.

- Agreed. We have recently published analyses (Doraisami et al., 2025, *Ecosystems*) from the same site, that evaluate the role that variation in wood carbon fractions play in deadwood carbon dynamics. These and related findings will be incorporated into a revised manuscript. Concepts of belowground biomass and wood carbon fractions will also be included. This content will be focused on discussing variation in wood carbon fractions of roots, in relation to belowground carbon accounting (Page 12, Lines 459-463).

4. Further optimization of charts and text. Some charts (such as Fig 2, 3, 4) are very information-dense, but some legends or labels (such as the correspondence between colors and taxa) are not prominent enough. Consider slightly increasing the contrast or adding clearer legend descriptions. When referencing charts in the text, refer more precisely to specific subplots (such as Fig 3A) to improve readability.

- Agreed. We will improve readability throughout our display items.

Biogeosciences Reviewer: 2

The topic is within scope. The paper addresses how wood CF assumptions influence forest C stock estimation, which is relevant to terrestrial C cycle research. The paper does not present new data, tools, or concepts, but provides a novel application of existing resources. The most original contribution is the comparative ranking of CF assumptions against species-specific data across scales, particularly the finding that the IPCC default performs worse than the simpler 50% assumption.

- Thank you for the review. We wish to note that this particular forest inventory dataset, and associated tree-level biomass and carbon estimation data, is indeed “new” as it has not been published in its entirety.

- The within-plot conclusions are well supported, and the ranking of CF assumptions is a useful contribution. However, the broader claims particularly related to the biome-scale extrapolation exceed what the data can support. The most substantial conclusions are operational rather than ecological: they inform which CF to use, not why CFs vary. The methods are clearly outlined, and the overall framework is sound. The tree- and subplot-scale results adequately support the main interpretations are strong.

- Good point. As noted in response to Reviewer 1, we wish to retain some content surrounding extrapolation, though we will temper these interpretations and reframe our extrapolation in a manner that motivates expanding our research to multiple forest sites (Page 10, Lines 393-397).

- The general approach is clearly described, but several details needed for full reproduction are missing:

1. The species-specific CF values used for Level 1 are not listed anywhere in the paper. A reader would need to go to the Doraisami et al. (2022) database, identify which entries correspond to the 30 HFDP species, and determine which value to use when multiple entries exist for a species. A supplementary table listing the CF value used for each species would make this reproducible.

- Good point. These will be added as a supplement.

2. It is not stated how species without a species-specific CF in the database were handled at Level 1. Were all 30 species covered? If not, what was the fallback?

- Noted. Our revised manuscript will be more specific in this explanation. Species-specific wood CF values were available for 17 of the 28 species in the dataset. Though these 17 species represent 99.9% of the total tree-level carbon stocks in our study site (based on our Ctree1 estimates). In the cases of the less common and small understory trees that did not have species-specific wood CF data, Ctree1 estimates are based on mean wood CFs from temperate angiosperms. This has been clarified in our revised manuscript (Page 5, Lines 150-154).

3. The method for aggregating subplots into 10 hectares (line 190) references "Doraisami et al., unpublished," which is not accessible. The reader cannot reproduce which subplots were assigned to which hectare. A brief description would help.

- Agreed. This noted citation is now published and cited in our manuscript. Also, note that the visual aggregation of subplots into 10 hectares is visually presented in Figure 4.

4. The code is described as available "upon request" (line 492) rather than deposited in a repository. This study contributes through its computational method and lack of codes is a missed opportunity. The analysis should be fully reproducible from the ForestGEO data and a short script.

- Agreed. Our code will be deposited in the University of Toronto Borealis Repository.

- Credit to prior work is generally adequate and the contribution is identifiable, and the novelty claims are precisely stated. However, the reference list is heavily weighted toward the authors' own prior publications and could benefit from broader engagement with related work by other groups.

- Good point. We will expand our citation list.

- The title is accurate and appropriately scoped. The abstract is complete but could be more concise. It would benefit from explicitly stating the counterintuitive IPCC finding and from qualifying the biome-scale extrapolation.

- Agreed.

- The overall structure of the paper is clear and logical. The presentation could be improved by leading the Discussion with the novel findings (they come later in lines 330,345) and trimming the speculative discussion of intraspecific variation, lines 384-404 (this is interesting but tangential. The paper has no data on intraspecific CF variation; it uses species means throughout. This section raises a question the study cannot address and doesn't connect back to the results). The language is clear and readable throughout. The writing is careful and precise with terminology. Non-specialist readers would be able to follow the logic.

- Good points. Novel findings have been elevated to the front of the Discussion section (Page 9, Lines 348-375). We have also reduced the content about intraspecific variation, while retaining some content related to Reviewer 1's request for content surrounding intraspecific wood CF variation and its role in forest carbon accounting (Page 12, Lines 455-463).

- The bigger concerns are related to statistical methods:

1. The ha-scale ANCOVA models (Table A3) fit 5 predictor terms to $n=10$. At this ratio of parameters to sample size, the high model R^2 values (0.889–0.999) are likely overfit. The subplot-scale models ($n=368$, Table A2) already show the same ecological patterns and have adequate replication. I would suggest simplifying the ha-scale analysis to 1-2 predictors or presenting the ha-level results descriptively and allowing the subplot-scale models to carry the inferential weight of the paper. Related to this, the biome-scale (794M ha) extrapolation in the Discussion (lines 336–342) would benefit from more explicit caveats about representativeness or could be framed as illustrative based on a much simpler model.

- Agreed. We appreciated the comment about presenting the ha-scale analysis and low sample size. In the revision we pair this back and interpret more cautiously (as noted above as well).

2. The paired t-tests (lines 165–170, Table 2) compare carbon estimates that are calculated from the same trees using different CF values. Because of this, the differences are built into the calculation rather than arising from independent variation in the data. With such a large sample size ($n = 39064$), even very small differences will produce extremely large t-statistics (e.g., $t = 526$), so the reported p-values are not informative about whether the differences are meaningful. In addition, the analysis treats each tree as an independent observation, which is not appropriate given that the trees are grouped within species and space: trees of the same species share CF values, and nearby trees or subplots are likely to be similar. This creates pseudoreplication and inflates statistical significance. A better approach would be to focus on the size of the differences (effect sizes) and how they change across spatial scales, or to use a mixed-effects model that accounts for species and spatial grouping (e.g., random effects for species and subplot). At the hectare scale ($n = 10$), statistical testing is likely underpowered, so results should be presented descriptively and inferences made with appropriate caution (see pt.1).

- **Agreed. We replaced our paired t-tests with a more robust random effects model, at the tree scale (Page 5, Lines 175-188). We have also noted caution in interpreting our per ha-sale results (Page 9, Lines 321-345). We find that the results remain primarily unchanged, though this is a statistically better approach.**

- The strength of this study lies in the complete census of 39064 trees across 13.5 ha, which allows CF-induced differences in C estimates to be calculated exactly at every scale without sampling uncertainty in the inventory. Several of the findings follow predictably from known CF values and allometric equations (e.g., that a 50% CF overestimates C, or that CF errors vary by species and tree size). The paper would become stronger from reframing its contribution: rather than presenting these results as new ecological findings, the authors could position the study as a definitive reference case — the first complete accounting of how CF assumptions propagate through a fully censused forest. The finding that the IPCC default (0.47) produces larger absolute errors than the 50% assumption (lines 330–334) is counterintuitive and could be given more prominence, along with the conditions (species composition, size structure) under which this holds.

- **Very much appreciated. We have revised our manuscript in a manner that highlights our contributions: 1) a first definitive reference noting the role wood CFs play in forest carbon accounting; and 2) our counter-intuitive finding that IPCC defaults may not necessarily be “better” than a 50% assumption (noting we only find this through analysis of a complete census) (Page 10, Lines 388-392).**

- Minor corrections:

1. Line 136: "βbrachnes1" should be "βbranches1" **Corrected.**
2. Line 157: "Ippc, 2006" should be "IPCC, 2006" **Corrected.**
3. Line 159: "HDFP" should be "HFDP". **Corrected.**
4. Line 435 (Fig. 1 caption): "Geogrphical" should be "Geographical" **Corrected.**

5. Line 436 (Fig. 1 caption): "contets" should be "contexts" **Corrected.**
6. Line 242: "median range of Csubplot values=4,278-4,580 Mg C 400 m⁻²" should be "kg C 400 m⁻²" **Wow. Eagle eye. Thanks. Corrected.**
7. Line 240: t-statistic subscript reads "t₃₉₀₆₃≥24.1" but this is a subplot comparison (n=368 subplots), not tree-level. **This has been removed as per comments about our t-tests.**
8. Line 196: Methods text says "Csubplot2 through Csubplot4, and Cha1 through Cha4", the second part should read "Cha2 through Cha4" **Corrected.**
9. Line 350: "(Fig. 3, 3)" **Corrected.**
10. Line 355: Extra closing bracket in "Doraisami et al., 2024))]" **Corrected.**
11. Table A2: The Csubplot1–Csubplot5 model shares identical parameter estimates with the Csubplot1–Csubplot4 model for the intercept, small tree biomass² and conifer proportion. **We have checked these values, and they are in fact correct.**
12. Table A3 caption: "Mg C subplot⁻¹" should be "Mg C ha⁻¹" **Corrected.**
13. The main text consistently uses "gymnosperm/angiosperm" but Tables A1–A3 and their captions use "conifer." Either make the terms consistent or add a note to explain why the terms are used interchangeably. **This has be streamlined to "conifer" throughout.**
14. Line 137-138: cites "Lambert et al., 2011" but the reference list only contains Lambert et al., 2005. **Noted. This has been changed to Lambert et al. 2005.**
15. Upon printing, the legends in Fig. 3-4 are hard to read, especially "Conifer proportion" values are impossible to read in both and "Carbon stock difference" values in Fig. 4. If possible, increase font size for all the legends and increase figure resolution. **Agreed. Figures 3 and 4 have been updated such that these legends have been made larger by 30%.**
16. The ± notation is used inconsistently: sometimes for standard error (line 69, "46.5±0.3% [s.e.]"), sometimes unclear whether it's SE or SD (line 39, "~870±61 Pg of C"). **All instances of error terms have now been clarified on their usage. Certain of these values (e.g., those in the first paragraph) denote 95% confidence intervals as reported in the initial papers, while others (e.g., those in our own results) denote different types of error terms as needed. No less, they are all individually clarified in our revised manuscript.**
17. Tables 1 and 2 could be combined into one table. They describe the same comparisons where one gives the absolute values and the other gives the differences. Merging them would reduce redundancy and let the reader see both side by side. **We tend to find that this complementary information in separate tables is preferable. While we agree**

that these could be merged, the resulting table would be massive and probably a bit confusing when discussed throughout the Results section.