

RESPONSE TO REVIEWER COMMENTS

Reviewer #1

General comment

I thoroughly enjoyed reading this paper. The dataset is extensive and precisely fills a long-standing gap in terrestrial records. This work will be highly impactful for researchers investigating deep-time terrestrial carbon cycles, hydrocarbon source rocks, and the interactions between life and climate on land. Its potential applications are broad, and I anticipate it will be widely utilized across diverse research fields. The manuscript is also clearly written and highly accessible, which is a significant strength. Only a few minor revisions are needed, all of which are straightforward to implement.

Thank you very much for your encouraging and thoughtful feedback. We are delighted to hear that you found the dataset valuable for filling the terrestrial record gap and that the manuscript is clear and accessible. We have carefully reviewed your suggested revisions and have addressed each point thoroughly in the revised manuscript..

Specific comments

First, the figures feel a bit rough and basic right now. They do the job, but for a journal like ESSD that aims for high visibility, they could look a lot nicer. A round of polishing (cleaner colors, sharper layout, better fonts and legends) would make the whole paper feel more professional and attractive.

Thank you for this suggestion. We agree that the figures would benefit from further refinement. In the revised manuscript, we have improved all figures by adopting a cleaner color scheme, sharper layouts, and more consistent fonts and legends, which we hope gives the paper a more professional and visually appealing presentation.

Second, age control is always the biggest challenge in terrestrial sediments. The authors explain their interpolation method clearly, but many readers will worry that it could create extra uncertainty or put samples in the wrong time bins. I'd strongly suggest adding a simple column (or at least a confidence level) for age uncertainty in the actual database file. It would also be good to add one short, honest paragraph in the Usage instructions section that directly tells people about this limitation and how to handle it. That small addition will make the dataset much safer and more useful.

We thank the reviewer for this important suggestion. While assigning a quantitative confidence level to each sample is not feasible given the nature of the data, we have addressed the spirit of this comment in two ways. First, we have added a new column "Age accuracy" to the database that clearly distinguishes between linearly interpolated ages (the majority of samples) and expert-estimated ages (a small subset). Second, we have added a short, candid paragraph to the Usage Instructions section that explicitly acknowledges the limitations of linear interpolation

where only two tie-points are available, and provides practical guidance on how users should account for this uncertainty—for example, by exercising caution in high-precision analyses. A dedicated discussion of these limitations and user recommendations has also been incorporated into Section 2.2 (Age field). We believe these additions make the dataset considerably more transparent and safer for reuse.

Third, the data availability part feels a little incomplete. The Geobiology Database link is mentioned, but actually going online to browse or filter the data isn't very convenient yet. The Excel file is easy to download, but keeping it updated over time could be tricky. I strongly recommend giving at least one additional, user-friendly and regularly updatable link (maybe a simple web dashboard or a clearly versioned live repository) so people can access and explore the data more smoothly and the authors can add new entries efficiently.

Thank you for this thoughtful suggestion. We would like to clarify that the Data Availability section already includes a Zenodo link in addition to the Geobiology Database and Excel file. Zenodo provides a stable, versioned, and publicly accessible repository that supports straightforward data download and accommodates future updates through new version releases under the same DOI. We believe this arrangement satisfactorily addresses the need for a user-friendly and regularly updatable access point, allowing both users and the authors to engage with the most current version of the dataset efficiently.

Minor points:

In the abstract (line 13), “exclusively” feels a bit too strong because quite a lot of data come from transitional facies and wood/coal samples: “primarily” would be safer and more accurate.

Thanks, we have changed “exclusively” to “primarily” in the abstract (line 13).

Line 19: “source-rock assessment” is better without the hyphen — just “source rock assessment”.

Thanks, we have removed the hyphen (line 19).

Please carefully check consistency of field/attribute names between the main text and Table 1 (e.g., line 85 SampleID vs Sample ID, line 87 Strata vs Stratigraphic Unit).

Thank you for catching this. We have carefully reviewed and standardized all field and attribute names to ensure full consistency between the main text and Table 1.

Around line 94 the text uses “PalaeoLat”, but Table 1 shows “Paleolat” — please pick one spelling and use it everywhere.

Thanks, we have standardized the spelling to “PaleoLat” throughout the manuscript and Table 1.

In the attached data table there is an obvious typo: “Yancahng FM.” should be “Yanchang FM.” —

worth a quick scan of the whole table to catch any others.

Thank you for spotting this. We have corrected "Yancahng FM." to "Yanchang FM." and conducted a thorough check of the entire text and data table to identify and correct any similar typographical errors.

Reviewer #2

Large-scale terrestrial plants colonized landscape since the Devonian, while the first angiosperm occurred in the Cretaceous. This work compiles terrestrial 49016 TOC and 17571 organic-carbon isotope data from Devonian to Cretaceous. These data cover the evolution processes of land plants

and its associated organic-carbon isotope. It will facilitate the future studies of long-term carbon cycles and terrestrial ecosystem evolution. I recommend it to publish in this journal after minor revision.

We sincerely thank the referee for the positive and encouraging comments, and for recognizing the value of this compilation in advancing the understanding of long-term terrestrial carbon cycles and land plant evolution. We have carefully addressed all the points raised below.

1. The authors mention a sparse to nearly absent data organic-carbon isotope data in Devonian and Carboniferous. However, such data are very important for the study on plant evolution on land. Please try to get more data.

We appreciate this important observation. We have conducted an extensive additional literature search; however, high-quality, well-dated terrestrial organic-carbon isotope records from the Devonian and Carboniferous remain genuinely scarce in the published literature. The paucity of data for these intervals reflects a real gap in the existing record rather than an oversight in our compilation. We have noted this limitation explicitly in the manuscript and hope that the present dataset will help motivate future fieldwork and geochemical analyses targeting these critical intervals in early land plant evolution.

2. What are the other terrestrial facies besides lacustrine, deltaic and river?

Thanks. In addition to lacustrine, deltaic, and fluvial settings, our dataset also encompasses alluvial fan, floodplain, swamp/peat mire, and coastal plain facies, providing a more comprehensive representation of the diverse terrestrial depositional environments documented across the Devonian to Cretaceous interval.

3. Adding a figure of organic-carbon isotopic compositions curve from Devonian to Cretaceous would be interesting.

We thank the reviewer for this constructive suggestion. We have added a new figure presenting the long-term trend of terrestrial organic carbon isotope compositions ($\delta^{13}\text{C}_{\text{org}}$) spanning the Devonian to Cretaceous, incorporating a moving average to better illustrate secular variations. We believe this addition substantially enhances the visual impact of the dataset and provides readers with an immediate sense of the broad temporal patterns captured in the compilation.

Response to Anonymous Referee #3 (RC3)

The authors present a database of 66,587 published TOC and carbon isotope measurements spanning 419-66 Ma, including robust metadata fields. The database will surely be of great benefit across many geoscience fields, both for current and future work.

The article itself is a short description that supports the publication of the data set. The data set is significant; it is unique, useful, and complete (see my one comment for Section 3.2), and the data set itself appears to be of high quality based on my own general screening and the description within.

Thank you very much for your positive and constructive feedback on our database and manuscript. We are glad that you recognize the value and uniqueness of this dataset for the geoscience community. Below we address each of your line comments point by point, and we have revised the manuscript accordingly.

Line comments:

37- this is dependent on depositional setting.

Thanks for pointing out the need to qualify this statement. The original sentence indeed implied a direct relationship between elevated TOC and paleoproductivity without sufficient attention to depositional setting. We have now revised the sentence as follows to explicitly acknowledge that the interpretation of TOC depends strongly on depositional conditions (e.g., redox regime, water restriction, sedimentation rate): "Because sedimentary organic matter is largely derived from aquatic primary producers (e.g., plankton and algae), elevated TOC values generally indicate enhanced paleoproductivity and increased organic carbon burial rates, though this relationship is strongly modulated by depositional setting (e.g., anoxia, restricted circulation, upwelling) and preservation conditions".

76- Database.,

We thank the reviewer for this careful observation. The erroneous punctuation ("Database.,") has been corrected to "Database," in the revised manuscript.

Fig 1. Facies is spelled wrong (Facecies). Data Sources is spelled wrong (Data Sourcies).

We thank the reviewer for spotting these typographical errors. "Facecies" and "Data Sourcies" have been corrected to "Facies" and "Data Sources" respectively in the revised figure.

146, 157- I'd caution use of academic theses and conference proceedings, which are not peer-reviewed. I suggest a filter for peer-reviewed vs not peer-reviewed datasets.

We appreciate this valid concern. We have added a "Peer-reviewed" flag (Yes/No) to the database, allowing users to filter out non peer reviewed entries as needed. A corresponding note has been included in Section 2.5 advising users to apply this filter for analyses requiring stricter source quality control. We also note that theses and conference proceedings constitute a small fraction of the total entries (less than 5%) and were retained for completeness; the new flag ensures full

transparency and gives users the flexibility to exclude them at their discretion.

Fig 3. Is this by publication? So 619 bubbles?

We thank the reviewer for raising this point. To clarify, each bubble in Figure 3 represents a unique section or borehole rather than a publication, the 619 figure refers to the number of publications in the database, while the total number of sections/boreholes is 774. To avoid any confusion, we have added a sentence to the figure caption explicitly stating: "Each bubble represents a unique section or borehole."

Section 3.2. There is a discussion on why the dataset cuts off at 419 Ma, but not as to why the database cuts off at 66 Ma.

Thank you. We have added the following explanation to the end of Section 3.2 to justify the upper boundary at 66 Ma. The Cretaceous - Paleogene boundary was chosen as the upper cutoff for several reasons. The Cenozoic has already been extensively studied, with numerous well-established, publicly available compilations covering terrestrial organic geochemical and isotopic records. Our primary objective is to address the comparatively under-explored Paleozoic and Mesozoic intervals in terms of integrated TOC and carbon isotope databases. Furthermore, the much larger volume of published Cenozoic studies would dramatically increase the number of entries and disproportionate curation efforts, while the wider array of high-resolution proxies available for the Cenozoic, but not for older periods, would introduce methodological heterogeneity that could compromise the internal consistency of the database. For these reasons, temporal coverage was deliberately restricted to 419 - 66 Ma.

214- Can you list somewhere what facies fall under "other"?

Thank you for raising this. We have made two changes to address the ambiguity. First, the adjective "other" in the sentence "Records from all other depositional facies are generally limited" has been replaced with "remaining" to avoid confusion with the database category named "Other." Second, we have added a sentence in the main text to clarify which facies fall under that category: "The 'Other' facies category includes paleosol, glacial, and other continental settings that are only described as 'continental' in the text."