COVER LETTER FOR NEW SUBMISSION

TITLE

Evaluation of 210Pb dating models using simulated datasets

[ 21/07/2021 ]

Attn: Brian Reich  
Editor-in-Chief  
Journal of Agricultural, Biological, and Environmental Statistics

REF. JABE-D-20-00258

Dear Professor Reich

We thank you for the quick response as well as the thoughtful and constructive comments on the previous version of this manuscript (JABE-D-20-00258) titled “A simulation study to compare 210Pb dating analyses”. These comments greatly contributed to improve the entire study. We recognize the need for both minor and some major revisions and have carefully considered each comment, from clarifying information concerning the statistical terminology to the inclusion of the remaining dating models in the analysis and experiments.

We feel that this paper will be of interest to the statistics community, as well as the paleoecological and environmental change communities.

The study clearly highlights the benefits of using Bayesian methods to provide reliable and consistent age-depth relationships by quantifying and evaluating model precision (bias) and accuracy (uncertainty). Plum is the first Bayesian model to date sediments using lead-210 (210Pb) and its use is growing with greater focus on recent (Anthropocene) changes in terrestrial (lake, peatland, wetland, soil) and marine environments. There are several studies submitted and in prep to paleo journals that are using Plum. 210Pb dating is now widely used in paleoecological, environmental, pollution and climate change studies to calculate rates of sediment and carbon accumulation or timing of change following environmental or land use change. As such, realistic estimates of age and uncertainty are critical as results can potentially have a high impact on policy-making and public perception.

We have outlined our responses to the comments on the previous version of the manuscript below for your consideration. We look forward to your final decision on this updated submission and thank you for your assistance in working towards publishing our work.

Sincerely,

Dr. Marco A. Aquino-López

Dr. Nicole K. Sanderson

Dr. Maarten Blaauw

Dr. Joan-Albert Sanchez-Cabeza

Dr. Ana Carolina Ruiz-Fernandez

Dr. J. Andrés Christen

\*\*\* Associate editor comments \*\*\*  
  
This manuscript reports the results of a simulation study to compare two methods (CRS and Plum) for determining ages from 210-Pb data. The simulation study appears to indicate reduced bias and improved coverage for Plum compared to CRS. Some discussion is provided to address why these results were observed.   
  
Although this work may be of interest to JABES readership, I cannot recommend publication in its current form. A resubmission addressing the following comments may be sent to reviewers.   
  
General comments:  
  
Models  
  
The manuscript lists these models used for 210-Pb data: CF:CS, CIC, CRS, CI-CRS, R-CRS, and Plum. In the primary analysis only CI-CRS and Plum are used while R-CRS is used in a follow-up analysis. An argument is made to focus on CRS based on the plurality of manuscripts using CRS and due to CRS being more flexible than CF:CS and CIC. This argument seems insufficient, partly because Plum is apparently infrequently used, and thus I suggest a resubmission include all methods for analysis of 210-Pb data.

The analysis was performed for every model, but the discussion focused on two methods (the new Plum Bayesian model and the classical CI-CRS model) as similar results were obtained. An Appendix was added to show the model comparison results from the other methods.

The only model that is provided in any mathematical detail is Plum. A resubmission should include mathematical detail for all models so that differences between the models can be understood. In addition, methods of estimation for these models should be detailed.

This was amended and additional references to the model applications were included  
  
Evaluation  
  
A collection of simulations were performed, but the structure of these simulations is unclear.

Table 1 provides different age-depth functions and tables in the Supplementary Material appear to show information about the simulations (although no caption is provided so it is difficult to tell). If possible, a single table should be provided to articulate the full set of simulations performed.

We expanded on the explanation on how the datasets were obtained and the GitHub repository was amended so it is publicly accessible.

The authors use the term "offset", but I believe they are talking about "bias". Standard statistical terminology should be used. In addition, the "normalized offset", where normalization refers to division by the standard error, is used to assess accuracy of methods, but standard evaluation criteria is coverage and bias

We thank the editor for this observation and all terminology was amended throughout the main text, figures and discussion to be consistent with statistical terms.

The main evaluation concerns CRS at low information levels. It is unclear why this should be the primary comparison of interest. Is it common to have low information levels? Why is there no discussion of the methods under the different functional forms? How about discussion of any of the parameters that were set/adjusted during the simulation study?

The functional forms were selected as they reflect what is most commonly observed in practice; these functions were chosen with the collaboration of our Palaeoecological co-authors.

It is also important to note that the age-depth function is the primary variable of inference, which is why we focus our discussion on it. While other variables such as the supported 210Pb are clearly important, their effect on the resulting chronology requires a separate discussion and set of experimental tools, which are beyond the scope of this study.

Specific comments:  
  
Figure 1 is better represented as a sentence.

We feel that the figure is necessary, as the original study cited did not focus on the popularity of the methods. The figure was a reinterpretation of the data of said study.  
  
"providing different chronologies even when the same model and dataset was used." - so what is different? calibration? but then isn't that a different model?

As it is mentioned in the text, user (lab) intervention, i.e. model choice and application, is the main driver of the difference between the results of these models. The text now makes more emphasis on this point

"hindrance variable"? perhaps nuisance? What does "supported 210-Pb" hinder?

The supported activity is a source of replenishable 210Pb in any sediment, and because the excess 210Pb cannot be directly measured (proxies like 226Ra exist but are not a direct measurement) and distinguished from the supported activity, this variable hinders the calculation of the excess 210Pb and therefore the age-depth model.

I think it would be useful to introduce an example data set, so that the reader knows what quantities are available for the analysis and can use this information to help interpret eq (1)

Data from a well-known core was added to explain the data itself and the models

as Plum is a new model, some of the details should be explained when eq (1) is introduced:  
- why use a normal distribution for a strictly positive quantity?  
- what is \bar{t}?, what is P\_i^S?  
- what are known and unknown? \Phi\_i, \lambda, \delta, \sigma\_i, \rho\_i?  
- is depth literally depth in the soil/rock/whatever? is this measured from the "surface" or from the "bottom"?  
- does it make sense for the age-depth function to be piece-wise linear? is this just a simplifying assumption?

The first and last point are part of a larger discussion and are covered in the cited work.

The rest of the points are now clarified in the main text.

The simulation error structure can be simplified by   
  
y \sim N(C\_{\hat{x}} + X\_{shift}, y^2\_{scat} + \sigma^2\_{R})   
  
with  
  
X\_{shift} \sim (1-p\_{out}) \delta\_0 + p\_{out} Unif(-x\_{shift}, x\_{shift})  
  
where \delta\_0 is a point-mass at 0.   
  
This alleviates the confusing notation of C\_{\hat{x}} -> \theta -> \theta' -> y(\theta'). I don't believe \mu(\theta') is defined, perhaps it is "This variable" at the top of page 7, but "This variable" refers to \sigma\_{min}.

The notation was changed to clarify this points.

It is not clear where 5333 simulations come from. A single table, similar to Table 1 that provides all the variables that were adjusted/et during the simulation study would be helpful.

The text was amended to clarify how the simulations were obtained and how they are used in the study.

Since Plum is a Bayesian analysis, shouldn't you state "credible interval" rather than "confidence interval", e.g. page 8, line 49. Has Plum been proven to have the appropriate coverage?

Text was amended to refer to credible interval when referring to Plum and confidence interval for any other method   
  
Submission seems rushed, e.g.   
- "nor improves does its accuracy improve", "of the accuracy a model", "both the effects of both the"  
- inconsistent notation 210Pb vs $^{210}$Pb vs $^{210}Pb$  
- references to Table 2.1 actually refer to Table 1  
- typos, e.g. "centred"  
  
  
Why not include coverage of the intervals? It seems like coverage is a more natural quantity to report than "offset" and indicating normalized offset being off by 2 or more indicates lack of coverage.

Should "offset" be called "bias" since "normalized offsets, presenting the distance between the modelled age and the true age normalized divided by the standard deviation". And should "standard deviation" here be "standard error"?

This comments were taking into account and new text refers to coverage and bias.

Supplementary material is not included. Provided github repository gives 404 error.

The GitHub repository was made publicly available so this problem should be fixed.   
  
Table at the end of the manuscript has no caption. Perhaps this is the expanded version of Table 1 that I was suggesting.

Tables were removed as in order to avoid confusion, the data can be found in the github repository, which is now publicly available.

No information about estimation was provided. How is the CRS model estimated? How is the Plum model estimated?

This is mentioned in the text (using the original equations and R packages)

References:  
- There are only two statistics references. One of these is in JABES from the same authors.   
- The 2020 JABES article by the same author is not included (<https://link.springer.com/article/10.1007/s13253-019-00374-2>).

The references mentioned were fixed, and statistical references are limited as the classical methods were published in non-statistical journals.