In this study, Pfalz et al. present a very useful piece of software (LANDO) that serves as a universal wrapper for simultaneously applying multiple age-depth software packages to age-depth data points. Considering that different age-depth modelling packages each have their own unique approach, it is generally a good idea that researchers consider multiple age-depth modelling software packages to further understand the various software/methods and to see how choice of method might affect interpretation and why. In this respect, LANDO is a very valuable contribution to the research landscape, because it centralises the running of separate software packages into one interface. It is great that the scripts have been made open access on Github. I also think it is great that the authors took the time to make an accessible piece of software with clear installation instructions for Mac and Windows. However, seeing as I use Linux, I could not follow these instructions for installation. The manuscript does provide a good overview of what LANDO does, so I simply review the scientific parts here, as opposed to the LANDO software itself. Finally, I'm also impressed that the authors managed to get *Undatable* running at speed in Octave!

In general, I think the manuscript is very suitable for the journal *Geochronology* and will be of great interest to the readership. In my opinion, some work and clarification can further improve certain parts of the manuscript, which I detail below.

**Main points**

The main issue with the manuscript as it currently stands, in my opinion, is not related to the software itself or how it is described, but how the manuscript uses the LANDO software in an exercise in interpreting the performance of the various age-depth modelling software packages. When comparing the different packages, the authors state:

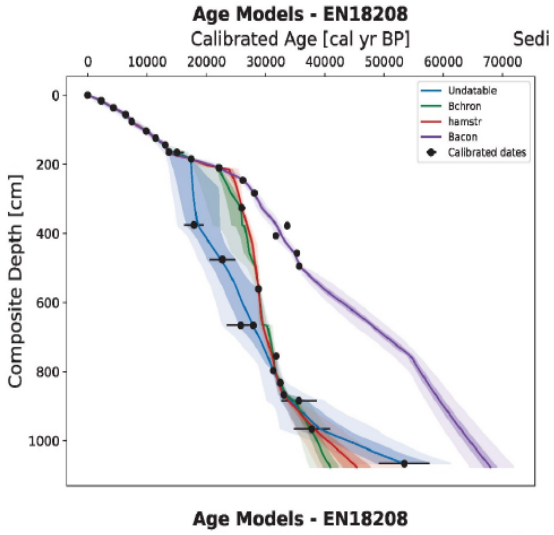
> *"To lower our impact and to avoid introducing biases in the modeling process, we used the default values from each modeling system as our own default values (Blaauw et al., 2021; Blaauw, 2021; Parnell et al., 2008; Dolman, 2021; Lougheed and Obrochta, 2019)."*

The above highlights the general issue with the parts of manuscript that compare age-depth modelling software packages. All of the age-depth model software packages in the manuscript are compared using "default" settings, but all of the packages have settings for a reason, namely that they should be adjusted. So it is possible that the age-depth software packages are not compared on their merits. I note that the LANDO software has the option to adjust the settings for each software package, so I am not describing a limitation of LANDO here.

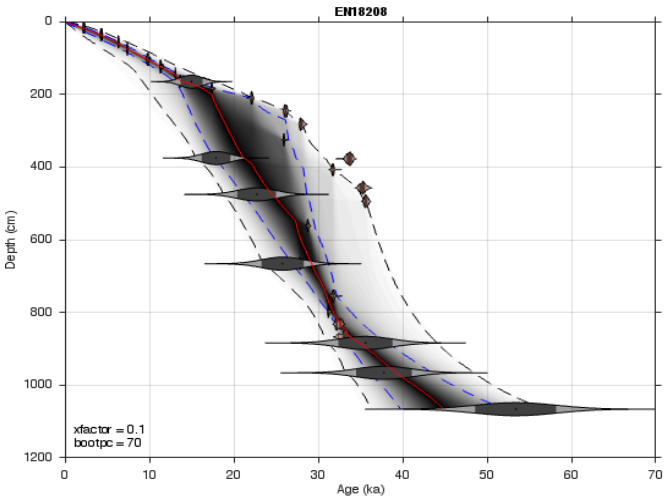
I can give an example about how using "default settings" can affect the interpretation in the case of *Undatable*. Figure 3 in the Pfalz et al. manuscript suggests that Undatable exclusively follows the younger dates between 200 and 600 cm, and the authors mention something similar in the manuscript in lines 408 to 410 of their manuscript.

While it is true that the GUI version of Undatable displays some settings in the data entry windows when the GUI first boots up, these are by no means "default values", but rather starting/dummy values in the GUI. The *Undatable* paper (Lougheed and Obrochta, 2019) discusses that bootpc (bootstrapping percentage) should be increased in the case of large age-depth scatter or age-reversals. Indeed, dealing with scatter in this way is stated in Lougheed and Obrochta (2019) as one of the main advantages of Undatable. Seeing as core EN18208 contains such scatter, I have rerun Undatable using a bootpc of 70 (after Gregor Pfalz kindly shared the input data with me), with the following result, with Pfalz et al. Figure 3 shown for comparison:

Pfalz et al. Figure 3




EN18208 with Undatable, 70% bootstrapping




In the above example, the Undatable uncertainty range expands to take into account the scatter of the dataset, and between 200 and 600 cm the highest probability area shifts more towards the centre of the age-depth scatter. This is the intended philosophy behind the deterministic *Undatable*, namely that the uncertainty range of the age-depth model should increase so that the scatter of the age-depth points is taken into account, i.e. 95% of the age-depth points should feasibly be located within the 95% uncertainty range of the age-depth model.

Other age-depth modelling packages also have their own settings and approaches.

## Other points

A small point regarding interpreting a lack of age-depth reversals as “undisturbed sediment”... Following bioturbation theory (e.g. Berger and Heath, 1968) when the sediment is fully uniformly mixed throughout the deposition history, downcore multispecimen / bulk samples will produce age-depth points that are in chronological order, i.e. lacking age-depth reverals. In other words, a lack of age-depth scatter is not an indicator for undisturbed sediment (despite perhaps 90+% of the literature assuming otherwise).

When describing the performance of the age-depth models, the text describes the age-depth models from top down, whereas most of the algorithms operate in the direction of sedimentation/time, i.e. from bottom up.

In the age-depth model figures, calibrated dates are indicated by black dots with error bars. Please add some information in the legend or caption about what the black dot is (median or mean calibrated age?), and the error bars (+/- 1sigma, i.e. symmetrical error bars, or the central 68% range, i.e. asymmetrical error bars).

The “optimised” age-depth model in Fig 4c takes what can be described as a middle route through the age-depth points, but with very small confidence intervals. It could be argued that such small confidence intervals mask the scatter of the age-depth determinations, and therefore the true geochronological uncertainty. This is more of a philosophical point, however, seeing as some age-depth packages try to find an optimised route between age-depth points with minimal age model uncertainty (e.g. Bacon, Bchron, OxCal), whereas others also expand uncertainty to take into account the scatter in age-depth points (e.g. Undatable). An argument can be made for either

approach, but in a manuscript that compares all the different types of approaches, it would be useful to point them out.

Thanks for providing this interesting manuscript to review, I look forward to seeing the finished product!

Kind regards,
Bryan Lougheed

P.S. Looking forward to seeing LANDO operating in the cloud (city).