Geochronology Discuss., author comment AC1
https://doi.org/10.5194/gchron-2021-39-AC1, 2021

Reply on RC1

Michael Dietze et al.

Author comment on "sandbox – creating and analysing synthetic sediment sections with R" by Michael Dietze et al., Geochronology Discuss., https://doi.org/10.5194/gchron-2021-39-AC1, 2021

Dear Pieter,

Thanks for your comments and careful thoughts. Let us quickly add some words to clarify and explain in this interactive discussion. Also, thanks for seeing some strengths in the package (incl. the bypassing of chart junk).

We made the application examples of 'sandbox' decisively crisp and easy to follow, especially the first two. The main goal is to let readers follow how sandbox works and how results can be extracted and further used. You are right, the questions of these two first examples could easily be solved with more simple direct approaches. Having said that, this is exactly the point that motivated us to contribute a generic framework for geoscientists. Instead of developing individual solutions for each arising question, one could also use a consistent method not requiring dedicated coding of math but instead the possibility to create a virtual twin - obviously more useful for purpose other than estimating the effect of sample container geometry on grain property distribution.

Regarding the third example, yes you are speculating correctly. We have imposed a larger age inheritance to larger grains. You may want to see this as the lower likelihood of large grains to see light during fluvial transport (a valid physical assumption that could even be written explicitly, see for example https://github.com/coffeemuggler/eseis/blob/master/R/model_bedload.R). We think this topic is picked up in the discussion (sec. 5.4) but you are right, it would help to point that out way earlier, when defining the 'sandbox' rules for this example.

We never argued that 'sandbox' would be process based, nor it would be a physical model. We see that the description of subpopulation definitions was not well delivered in the manuscript, though. It is true that one **can** assign size, density, and further parameters independent from each other. However, this is just the allowed flexibility of the framework. In other words, nobody would hinder you to tightly connect grain-size distribution and specific density if you want to let them covary, following for example a defined physical law. I would be happy to read your suggestions whether we should expand on that topic in the revised version of the manuscript. Our intention was to keep things simple, but you may be right, it would help to show how one can impose covariance among grain properties.

Actually, you could envision ‘sandbox’ as a tool that allows knitting physical process models into descriptions of depositional records. Adding physically based relationships among parameters could indeed be simply added (if wanted). All one needs to touch are the rules, i.e., the depth dependent parameter descriptions. If, for example you want to let modal grain-size change with over a million-year time series of average bed shear stress, then this could be easily propagated into ‘sandbox’. Likewise, grain-size scatter could be defined by time series of river "flashiness" or any other meaningful independent variable you want. Here the question is again, which of the examples should be meaningfully added to the manuscript to not inflate it to a "long list of possible things" (potentially leading to confusion). We are happy to read suggestions from the community.

We find the quartz vs. zircon example an intriguing one. Actually, ‘sandbox’ could do a pretty good job in simulating this effect - and to explore its consequences on sample composition or other properties written into the sampled grains. To include this phenomenon, one would need to define two populations: quartz and zircon, each with a depth-dependent size and density rule and a mixing proportion rule for the two populations. So, sure this additional example could become part of the manuscript. We are confident that there will be further such valuable questions to ask, and effects to test. This brings us back to first paragraph about why using sandbox instead of a dedicated analytical solution: this one generic framework allows testing this and many more such questions consistently with the "same approach", along with subsequent analysis chains. The effect of changing zircon proportions in sediment layers would become important for micro-dosimetry issues, but this will require true coordinate-based relationships, which is decisively not implemented in ‘sandbox’ and would be an overkill (there are also limits of the usefulness of R). Software to simulate particle-matter interaction, like *GEANT4* would be more advisable here.

We decisively picked Geochronology because the overarching theme of ‘sandbox’ is age-depth-parameter relationships of sedimentary deposits. This is a topic of paramount relevance for dating applications but often not exploited with the needed care. This becomes even more true, if the systematic effects of the grain size structure are considered as systematic uncertainties that go into, for instance, Bayesian modelling frameworks, such as BayLum.

The only true alternative (excluding journals from the non-open access and CC-BY realm) would be GMD. However, we are afraid that the circle of people we want to reach and inspire by ‘sandbox’ will not enthusiastically follow GMD as outlet of papers. When adding parameters like major or minor elements, cosmogenic nuclide concentrations as well as carbon fractions or other biomarker content of grain populations, we believe that ‘sandbox’ may well become an interesting tool for a wide range of scientists occupied with other than luminescence-dating routines. Here again, it becomes difficult to find a balance between providing illustrative easy-to-follow examples and expanding on what would be possible (and that is a lot, as the text above indicates). Perhaps one way to go is to start simple and with more and more studies using ‘sandbox’ as a tool, the wider scope of possible applications will emerge. We consider this contribution here as an introduction of the package. Further, more complex and maybe even more relevant state-of-the-art research questions would go beyond the scope here. However, they may fit well into separate studies, where individual problems can be explored in detail and accordingly, the package and functions may be expanded or adjusted to the new question under consideration. Overall, we are happy to see your suggestions and would be open to explore together how those could be investigated using sandbox

About the second "Other comment", we shall underline that parameters (and rules) can be defined using not just normal distribution functions, but also uniform, normal, gamma and exact, as described in the manuscript and package documentation. We initially also had the option to use empirical distribution functions, as well. However, the overhead and parameterisation hassle was significant, so we decided (for now) to skip this additional option, assuming that gamma, normal, uniform and exact (no scatter) already provide quite a bit of freedom to define parameters. We should point out though that the grain-size definition indeed is in phi scale. This limits flexibility of this, but only this one, parameter. However, that constraint was decisively made, in agreement with long standing support from sedimentology.

Last, we see that a) the manuscript needs to become clearer here and b) that ‘sandbox’ is not at the top end of possibilities.

Happy to follow up a further discussion,

Michael Dietze, on behalf of the co-authors