Geochronology Discuss.,
https://doi.org/10.5194/gchron-2020-6-AC2, 2020

Interactive comment on "Technical note: A prototype transparent-middle-layer data management and analysis infrastructure for cosmogenic-nuclide exposure dating" *by* Greg Balco

Greg Balco

balcs@bgc.org

Received and published: 27 May 2020

This responds to the review comments by Richard Jones. Overall, this review is generally supportive of the paper and agrees that the transparent-middle-layer concept highlighted in the paper is a potentially useful contribution to geochronology. It then goes on to discuss a number of points that focus on aspects of the prototype implementation of this concept.

I very much appreciate Dr. Jones's close attention to these areas in writing the review.

These points are all valuable and worthy of discussion, and I discuss them below at some length, but they are also somewhat ancillary to the main point of the paper, which is to discuss the overall transparent-middle-layer concept. Thus, at the end of this response I will try to return the focus to what the implications of these areas of discussion are for actually revising and improving the paper.

Basically, the bulk of the review includes discussion of two main areas: (i) some aspects of the online exposure age calculator that forms the middle layer in the prototype implementation, and (ii) some suggestions for improving the web server application that forms the analysis layer of the prototype application.

The first of these mainly focuses on the fact that the online exposure age calculator used in the prototype (version 3 of the online exposure age calculators originally described by Balco and others (2008) and subsequently updated) is not very well documented. This is true and this comment certainly has merit. The original online exposure age calculator from 2009 set a high bar for documentation: not only was a paper published describing it, but there was probably a hundred pages of additional online documentation of mathematical formulae, details of numerical implementations, and comprehensive instructions for using the MATLAB code. The latest version, on the other hand, only has fairly general descriptions of basic concepts, and is nowhere near as well documented as the original. The fact is that as the current version is more or less a purely volunteer effort at this point, it has not been feasible to get to the same standard. There are several potential solutions to this problem, the simplest one most likely being to establish a documentation wiki that can be contributed to by multiple users of the calculator, so as to spread the responsibility more widely among interested and engaged users such as Dr. Jones. Of course, this has not been implemented either, but at least there is a fairly clear path for how to do so.

From the perspective of the paper being reviewed, however, the shortcomings of the documentation of the online exposure age calculator used in the middle-layer calculations are somewhat beside the point. The whole point of the distributed cloud infras-

tructure used for the prototype implementation described in the paper is that there is no reason that there has to be only one middle-layer server: in principle, any of the existing online exposure age calculators could fill this role. As a practical matter, however, right now there is only one that both (i) is fast enough for dynamic exposure age calculations and also (ii) has a programmatic interface that allows other software to easily submit input data and receive clearly formatted results. The other two options are the "CRONUSCalc" online calculator of Marrero and others (2016) and the CREp system of Martin et al. (2017). CRONUSCalc does not offer a programmatic API, is currently too slow for convenient dynamic calculations, and returns results by email instead of http or another more usual protocol. CREp also lacks a simple API, requiring instead a file upload for data submission. If these online calculators were to develop a simple http- or other standardized-protocol-based API in future, they could easily be used as middle-layer elements.

The second main point of the review focuses on a number of suggestions for improvements to the ICE-D web server that forms the visualization and analysis layer of the prototype implementation. First, Dr. Jones correctly points out that it would be extremely useful if the ICE-D web server duplicated the feature of the online calculators that non-default production rate calibrations could be propagated into the results. I agree completely, and this would be reasonably straightforward to develop, but I haven't done it yet. The reviewer then goes on to highlight a number of other potential functions of the ICE-D web server that would improve his workflow for data analysis, including presentation of data in various aggregated tablular forms as well as improvements to the HTML coding to facilitate copy-and-paste operations between browsers and spreadsheets.

These are all helpful and valuable suggestions, but again focusing on the present paper, two things are important here. First, the purpose of the paper is to highlight the overall usefulness of a transparent-middle-layer architecture by describing a prototype implementation that has a number of features designed to show proof of concept and

also to be a guide for what could be developed in future. Clearly this is working, because the simple capabilities that exist in the prototype have inspired this reviewer to come up with ideas for lots of other more advanced or more specific capabilities. Second, in the same way that many different online exposure age calculators could occupy the middle layer, the whole point of the distributed architecture is that there can also be many different software applications occupying the visualization and analysis layer. Dr. Jones could develop, or cause to be developed, his own analysis layer application with features designed for his specific workflow. Of course, in the current prototype situation, that is a bit difficult because the various APIs are mostly not fully documented, and it is still necessary to interact with me personally to get access permissions to the various server elements, but I would be very happy to help achieve it.

Finally, to refocus the discussion on improvements or modifications to the present paper that are indicated by this review, my take on this is that the main needed improvement is to include additional discussion and emphasis of the principle that although certain example middle- and analysis-layer elements are used in the prototype, the whole point of the distributed architecture is that it makes the transparent-middle-layer concept agnostic with respect to what elements are used. The specific middle- and analysis-layer elements used in the prototype are not themselves an inherent part of the transparent-middle-layer architecture. Multiple calculation methods could occupy the middle layer, and any number of visualization and analysis applications can interact with the data and middle layer elements. In fact, this is the goal: no one analysis application is likely to meet the workflow needs of all users, and we shouldn't expect it to. The distributed infrastructure makes it possible for many users with different analysis or visualization needs to utilize common lower-level resources.

Interactive comment on Geochronology Discuss., https://doi.org/10.5194/gchron-2020-6, 2020.