

Interactive comment on “Arctic Holocene proxy climate database – new approaches to assessing geochronological accuracy and encoding climate variables” by H. S. Sundqvist et al.

Anonymous Referee #2

Received and published: 31 March 2014

This paper describes a major new compilation of Arctic paleoclimate records spanning the last 6,000 years. This new database is significantly more comprehensive than the collection of Arctic paleoclimate records that is currently publicly available from international data centers. The authors have also given much attention towards raising the standards of metadata associated with these records, for example adding fields to describe the inferred climate variable and the geochronologic accuracy. Most of my major comments below focus on these new metadata and how to best represent them, since they are so cutting edge. If these questions are addressed, it is possible that this paper will be frequently referenced for helping to establish best practices in metadata description in these sorts of compilations.

C160

Major comments

1. For each calibrated variable, please provide somewhere in the metadata the reported size of the calibration error. This is hugely important for comparing to model output, a stated goal for this compilation. A one paragraph discussion in the text of generally how large these calibration errors are would also be useful.

2. Issues with the Geochronology accuracy score: While I fully appreciate the attempt to provide an objective and quantitative score, there are some aspects of the scoring that seem illogical and overly complex.

a) For the uniformity of trend, I can't follow the definition in Appendix A. It is the root mean square error of individual dates with respect to the spline? Or the root mean square error of a linear trend with respect to the spline? The name “uniformity of trend” suggests to me that records with more linear trends should have higher uniformity scores. Is this the case?

b) For the reliability score p , it seems strange that having one or more rejected dates would lower the reliability. For example, sometimes researchers date a material in the hopes of getting reliable dates from it, determine that the dates for that particular material are not reliable, and reject them all. Why should such a data collection approach yield a lower reliability measure than measuring dates from just one material and not rejecting any of them?

c) In terms of age reversals negatively impacting the p score, it seems like having a high sampling density for ^{14}C (especially in the vicinity of a radiocarbon plateau) would lower the reliability score. Generally, though, high sampling density should improve reliability, not hurt it.

d) The accuracy score defined here seems like too much of a “kitchen sink” approach and might be more robustly defined using just 2-3 of the most important elements identified by the authors (eg, frequency, regularity, material). A different way of doing

C161

the weighting factors that might be less arbitrary is to average the percentile ranks for each of the elements being scored. The final accuracy score would, of course, just be relative to the other records in the data set (no improvement over current approach) but the score would have a physical meaning and not be on some arbitrary scale.

3. Formatting issues in excel file: Different names for the same variable are used across records, eg, upper depth, depth_up, depth top, top depth. There is also no consistent column ordering for the data or chronology tables (eg, depth, age, variable in that order). These formatting inconsistencies make machine reading very difficult. I wasn't able to locate the machine-readable vPIR text files from the WDC-Paleoclimatology site, so was unable to evaluate these.

4. It would be helpful to show a figure with the number of records available through time. I know that every record has data between 2-6 ka, but how quickly does the number of records drop off beyond 6 ka?

5. Figure 1: More information is needed to explain this analysis. Was the Reanalysis used for both summer (which months?) and the Arctic Oscillation? How is the Arctic Oscillation defined? What years were analyzed? Please provide citation information for the Reanalysis and credit to the data center that provided the Reanalysis output. Also, explain in the main text whether or not there is any indication that this same correlation pattern existed over the last 6k.

Minor comments

Page 8, Line 9: Equation 1 doesn't fully describe "records with an average sample resolution of at least 400 yr and two standard deviations of less than +/- 200 yr." Equation 1 can be simplified to: $R = (t1 - tn)/n$, which is just the average sample resolution. How is the standard deviation criterion expressed mathematically? Is it the same as the expression for regularity provided in the appendix?

Page 24, end of line 11: change "spine" to "spline"

C162

Table 2: Do question marks in the statistical detail column mean "unknown"? Please explain this abbreviation.

Please provide complete citations (not just eg, Smith et al. 2001) in the excel/archived files. This is essential to link back to the original publications and to give people full credit for their work.

Interactive comment on Clim. Past Discuss., 10, 1, 2014.

C163