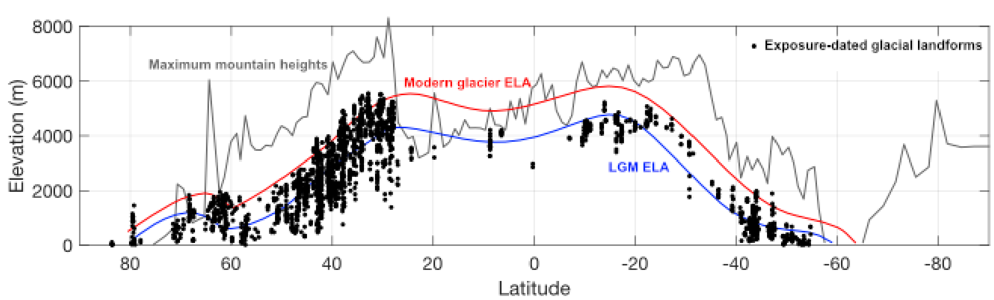
SYNOPTIC ANALYSIS OF GLOBALLY-DISTRIBUTED DATA SETS OF COSMOGENIC-NUCLIDE EXPOSURE AGES

Greg Balco, Berkeley Geochronology Center, Berkeley CA USA: balcs@bgc.org

EGU21-3513; https://doi.org/10.5194/egusphere-egu21-3513

This presentation describes a project to make large data sets of cosmogenic-nuclide measurements available for synoptic global analysis of paleoclimate, glacier change, and landscape change. It is based on the 'ICE-D' (“Informal Cosmogenic-nuclide Exposure-age Database”) transparent-middle-layer infrastructure for storing cosmogenic-nuclide measurements and generating internally consistent exposure-age data (Balco, 2020a). One part of the ICE-D project is the ICE-D:ALPINE database (http://alpine.ice-d.org), which includes more than 9000 cosmogenic-nuclide exposure ages from alpine glacial deposits, mostly moraines, in mountain ranges worldwide. Because glacier extent is a climate proxy, synoptic analysis of these data is potentially useful for paleoclimate analysis and model validation. The aim of the ICE-D project is to address a number of messy data-management and analysis problems associated with cosmogenic-nuclide data and their application to dating glacial landforms, thus making it possible to easily apply quantitative analysis to the entire globally-distributed data set.

Figure 1. Left, geographic distribution of 2562 exposure-dated alpine glacial landforms in the ICE-D:ALPINE database. Below, same data plotted in latitude-elevation space with approximate modern and LGM glacier equilibrium line altitudes (ELAs).



To explore this, this presentation focuses on how the data management system can be applied to test an example paleoclimate hypothesis related to a focus of research in glacial chronology in recent decades, specifically the geographic distribution of glacier response to cooling during the so-called Younger Dryas (YD) climate event between 11,500-12,700 years BP. It will highlight: (i) the difference between inferences based on single instances, which are common in existing literature, and error-tolerant hypothesis testing based on a large data set; (ii) means of quantifying the importance of the details of cosmogenic-nuclide production-rate calculations to testing paleoclimate hypotheses, and (iii) likewise, approaches to understanding the importance of geomorphic processes and landform evolution to global paleoclimate inferences drawn from exposure-dated landforms.

Example hypothesis. The example paleoclimate hypothesis is as follows: “There are no alpine glacier moraines of Younger Dryas age in the American cordillera.” Here “American cordillera” means the mountains of western North and South America between Alaska and Patagonia. Of course, the western cordillera is the location of all alpine glacial deposits in the Americas unless Greenland is included in North America, so I am using this term instead of the more general “North and South America” mainly to make clear that Greenland is not included. As noted above, the geographic distribution of glacier advances during the YD has been a focus of exposure-dating of glacial deposits in the mountains of the Americas for decades, ranging from Gosse and others (1995), who argued for YD cooling in western North America based on an exposure-dated moraine in Wyoming, through many similar studies, to Young et al. (2019), who argued for YD cooling in western North America based on an exposure-dated moraine in Alaska. Various other studies, which will not be reviewed in detail here, have concluded that YD moraines are or are not present in various other mountain ranges of North and South America. An important element of these two studies and many others (although not all) is that they are based on the logical principle that the identification of a single glacial deposit of YD age shows that the YD was expressed as a regional climate event. This approach is not error-tolerant: because only one deposit is being dated, it takes only one error in dating a deposit to lead to an incorrect result.

Prediction from the example hypothesis. Now, illustrate this hypothesis by considering a large number of hypothetical moraines in North and South America. Suppose that moraine emplacement following glacier advance occurred prior to the YD (e.g., during the Antarctic Cold Reversal) and after the YD (during the Preboreal/early Holocene), but not during the YD. Further assuming a uniform likelihood of moraine formation during the periods in which moraine formation is allowed, these assumptions imply that the histogram of the *true* ages of alpine glacier moraines in the Americas looks like Figure 2:

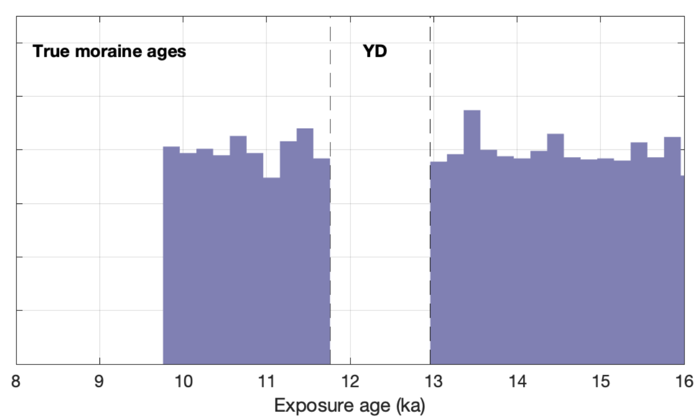


*Figure 2. Histogram of the expected distribution of **true** moraine ages if moraines were emplaced before and after, but not during, the Younger Dryas climate interval. This was generated by assuming that moraine emplacement is random and uniformly distributed in time during periods in which it is allowed. Small variations around the mean rate are due to the random nature of the simulation.*

However, exposure-dating of moraines is not arbitrarily precise and involves a variety of uncertainties associated with measurement of cosmogenic-nuclide concentrations and estimates of production rates. Thus, observed exposure ages of moraines differ from their true ages. If we assume that the dating uncertainty for each moraine is normally distributed with a standard deviation of 4% of the age (which is probably a minimum limiting value; see discussion in, e.g., Balco, 2020b), the distribution of *true* moraine ages shown in Figure 2 implies a distribution of *observed* exposure ages that looks like Figure 3:

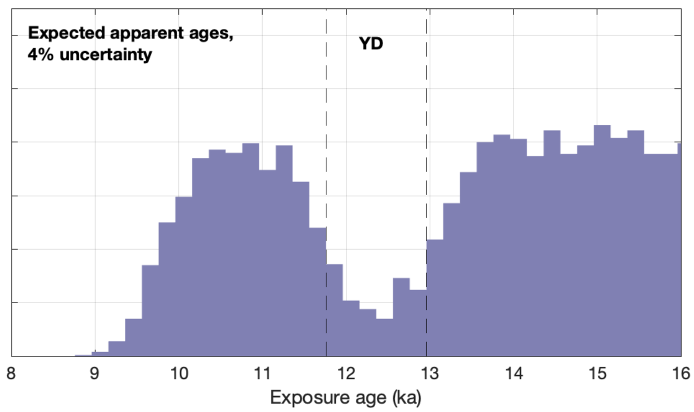


Figure 3. Histogram of the expected distribution of ***observed*** *exposure ages for moraines, if the distribution of true ages is as shown in Figure 2 and the dating uncertainty on each moraine is normally distributed with a 4% standard deviation.*

The point here is that if we start with the hypothesis that *zero* moraines in the Americas were emplaced during the YD, it leads to a prediction that, even so, we will observe *more than zero* moraines in the Americas with apparent YD exposure ages. This is because the duration of the YD is relatively short compared to even a best-case estimate of the uncertainty in exposure ages for moraines. This highlights the fact that exposure-dating of single moraines is not an error-tolerant way to test the hypothesis.

Comparison of prediction with observations. An alternative, error-tolerant, method for testing the hypothesis would be to compare the distribution of moraine exposure ages *predicted* by the hypothesis with an *observed* distribution of a large number of moraine exposure ages. This is possible by making use of the ICE-D:ALPINE database, which contains data for 813 moraines from the American cordillera. Applying an automated algorithm for outlier detection, averaging of multiple remaining ages from each moraine, and rejection of sets of ages that are unlikely to belong to a single population, the distribution of moraine exposure ages in the Americas generated from the ICE-D database is shown in Fig. 4:

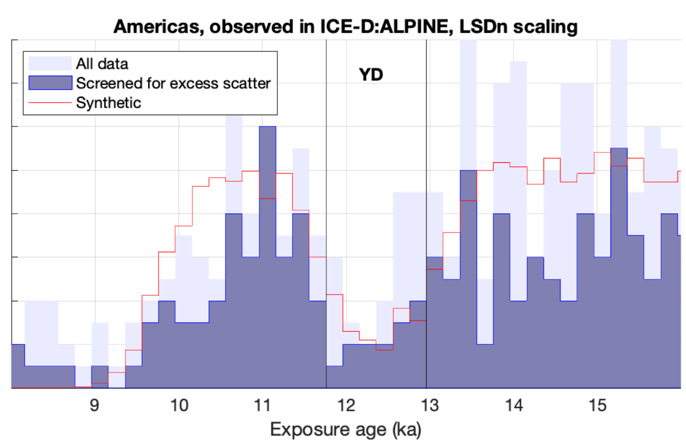


Figure 4. Histogram of the observed distribution of moraine exposure ages in the American cordillera in the ICE-D:ALPINE database. All data were subjected to an automated outlier-rejection and averaging algorithm. The difference between the light and dark blue histograms is that the data plotted in dark blue have also passed an additional screening test such that the hypothesis that remaining ages from a moraine belong to a single population cannot be rejected at 95% confidence. The red line is the synthetic distribution from Fig. 3.

This observed distribution strongly resembles the distribution predicted by the hypothesis that there are no YD moraines in the Americas: both display a large number of apparent moraine ages predating and postdating the YD, with a significantly reduced but not zero number of apparent moraine ages within the YD. Qualitatively, it appears that the observations are consistent with the hypothesis, which could lead to the conclusion that, in fact, the YD climate event was not expressed by glacier changes in the American cordillera. It is likely impossible on the basis of these data to *disprove* the hypothesis that exactly zero moraines were emplaced during the YD in the Americas. One could further apply various statistical methods[1] to quantify the likelihood that the two distributions are the same, and also to explore to what extent the likelihood depends on things like the uncertainty model used in generating the predicted distribution.

This result highlights the point that comparing observed distributions of large numbers of exposure ages with predicted distributions that take expected errors into account is an error-tolerant method of evaluating a paleoclimate hypothesis using exposure-age data. Even if many of the individual exposure-dating studies that make up the observed distribution include methodological, analytical, or interpretive errors -- which is almost certainly the case -- the conclusion will not be fatally affected by these errors as long as the likely magnitude and distribution of errors used in forming the predicted distribution is reasonable. The result also highlights that both simulated and observed apparent exposure age distributions contain many apparent ages within the YD: considering all ages between 9-18 ka, 7% of the apparent ages in the synthetic distribution and 9% of the ages in the observed distribution lie within the YD. A conclusion that YD cooling drove glacier change in the Americas that was based on one of these moraines would be incorrect. Thus, an error-tolerant approach based on a large data set that contains errors is most likely preferable to a non-error-tolerant approach based on single dated landforms, no matter how carefully chosen.

The effect of production rate scaling assumptions on hypothesis testing. Production rate calculations for exposure-age dating are sometimes quite complex, and a major challenge in using exposure-dated glacial deposits as a paleoclimate record is understanding how changes in scaling assumptions propagate through a large set of geographically widespread exposure ages and subsequent data reduction steps into a conclusion about paleoclimate. The transparent-middle-layer architecture of the ICE-D:ALPINE database is effective for investigating this. For example, Figure 5 replicates Figure 4 for three different commonly used production rate scaling methods ("St", "Lm", and "LSDn" in version 3 of the online exposure age calculator described by Balco et al., 2008 and subsequently updated).

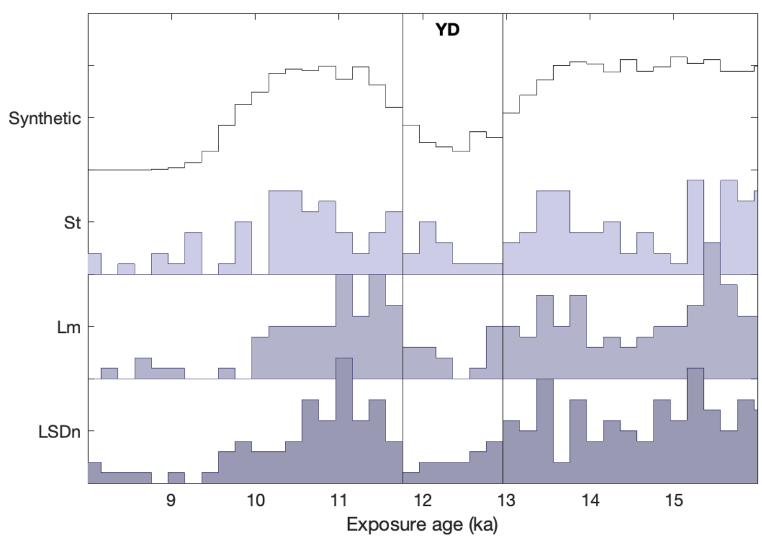


Figure 5. Histogram of the observed distribution of moraine exposure ages in North and South America in the ICE-D:ALPINE database,as shown in Figure 4 (darker histogram, including the test for excess scatter), but calculated with three different production rate scaling methods. The 'LSDn' histogram is the same as the one shown in Fig. 4.

Although the observed moraine age distributions display frequency minima near the YD for all scaling methods, the position of the minimum within the YD varies among the scaling methods. Thus, if there do exist moraines whose true age falls within the YD in the Americas, the St scaling method implies that they are more likely to be late in the YD, whereas the LSDn scaling method implies that they are more likely to be early in the YD. Thus, different production rate scaling methods might lead to a different conclusion about paleoclimate and glacier change in this example.

The effect of landform evolution assumptions on hypothesis testing. Like production rate scaling assumptions, assumptions about how postdepositional evolution of glacial landforms affects the distribution of exposure ages on the landforms can also propagate through the process of exposure-dating of glacial landforms into a paleoclimate conclusion. Individual exposure ages from a moraine are generally believed to differ from the true emplacement age of the moraine for three reasons: (i) analytical uncertainty; (ii) recycling of clasts that were exposed in a different location prior to moraine emplacement, leading to an inherited nuclide concentration and a spuriously old exposure age; and (iii) postdepositional disturbance leading to a spuriously young exposure age. As often discussed by, among others, Putkonen and Swanson (2003), Applegate and others (2010, 2012), and Balco (2020b), and as shown in Figure 6, these processes are expected to lead to a distribution of exposure ages on a moraine that is skewed young due to postdepositional disturbance and also has a long tail of outliers on the old side due to inheritance. Even if the old outliers are successfully identified and removed, central measures of the remaining distribution such as the mean or median are expected to underestimate the true age of the moraine due to the skewness of the distribution to the young side.

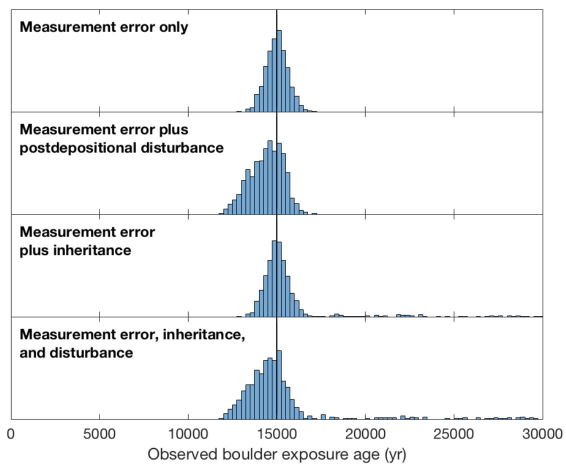


Figure 6. Effect of measurement uncertainty, nuclide inheritance, and postdepositional disturbance on a population of observed exposure ages on a moraine. The vertical black line is the true emplacement age of the moraine. Adapted from Balco (2020b).

The algorithm used in generating the histograms of observed moraine exposure ages for Figs. 4 and 5 is designed to exclude scattered outliers, so should theoretically account for nuclide inheritance. However, it assumes that ages remaining after outlier removal form a symmetrical distribution and then uses the mean of this distribution as the moraine age, so it does not account for postdepositional disturbance. A simple modification to the algorithm that could account for skewness due to postdepositional disturbance would be to use the maximum age, instead of the mean age, from the central distribution remaining after outlier removal. The effect of this on the example hypothesis test is as follows:

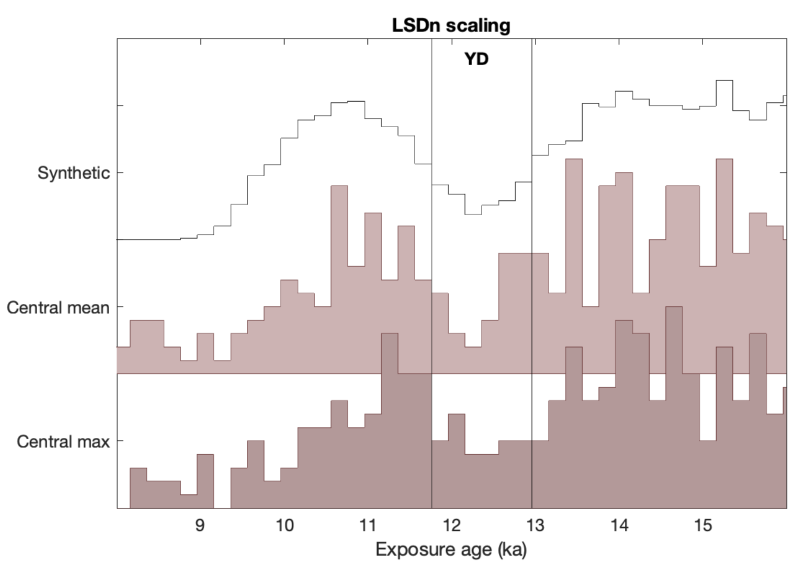


Figure 7. Effect of the simple algorithm to account for postdepositional disturbance on the test of the hypothesis that there are no moraines of YD age in the American cordillera. The histogram labeled “central mean” is the same as the “all data” histogram in Fig. 4. The histogram labeled “central max” uses the oldest exposure age remaining after outlier removal as the best estimate of the moraine age.

Using a maximum-of-central-distribution algorithm rather than a mean-of-central-distribution algorithm may improve agreement between the observed distribution of moraine exposure ages and the synthetic distribution predicted by the hypothesis that there are no YD moraines in the American cordillera; the period of anomalously infrequent moraines is wider and extends across the entire duration of the YD. The point of this is that a process-aware approach to aggregating multiple ages from the same moraine may lead to a different paleoclimate conclusion than a non-process-aware averaging scheme.

Summary. The main point of this presentation is that the amount of data available on past glacier changes from exposure-dating of glacial landforms is large enough, and the computational tools for dealing with it are good enough, that it is possible to move away from testing paleoclimate hypotheses using non-error-tolerant methods based on single carefully chosen landforms and toward error-tolerant methods based on large data sets.

In addition, a bigger-than-one-data approach using the ICE-D infrastructure is effective for fully propagating the effect of both production rate scaling assumptions and assumptions about geological processes leading to complex distributions of moraine ages through a fairly complex data analysis. However, the observation that both of these issues can affect the results of using a large data set of exposure-dated moraines to test a paleoclimate hypothesis highlights the need for an independent method to benchmark and validate both production rate scaling methods and methods for dealing with the effect of geologic and geomorphic processes on exposure-age distributions. For production rate scaling, a large calibration data set of exposure ages from landforms of known age can be used for this purpose. However, although there has been extensive work on geomorphic process modeling of exposure-age distributions (e.g., Applegate et al., 2010, 2012), there is no equivalent benchmarking data set for testing these approaches. At the moment, this is probably the major obstacle to using the global data set of exposure-dated moraines for paleoclimate reconstruction. It's worth thinking about ways to address this, perhaps by collecting large exposure-age and geomorphic process data sets from several independently dated moraines.

Acknowledgements

The ICE-D:ALPINE database would not have been possible without the contributions of Pierre-Henri Blard, Shaun Eaves, Jakob Heyman, Alan Hidy, Maggie Jackson, Ben Laabs, Jennifer Lamp, Sourav Saha, Mehmet Akif Sarikaya, Irene Schimmelpfennig, Perry Spector, Joe Tulenko, and Jill Pelto.

Endnotes

[1] Clearly it is disappointing to the reader to be left hanging without the results of, e.g., Kolmogorov-Smirnov or Anderson-Darling tests, but the author ran out of time before the meeting in preparing this presentation. Of course, the reference distribution that assumes an equal likelihood of moraine emplacement at all times before 12.9 ka is arbitrary, so it's not clear you would want to believe the result anyway -- if you get to totally invent the reference distribution, it's easy to make sure the observed distribution matches it.

References cited

Applegate, P.J., Urban, N.M., Laabs, B.J.C., Keller, K. and Alley, R.B., 2010. Modeling the statistical distributions of cosmogenic exposure dates from moraines. *Geoscientific Model Development*, *3*(1), pp. 293-307.

Applegate, P.J., Urban, N.M., Keller, K., Lowell, T.V., Laabs, B.J., Kelly, M.A. and Alley, R.B., 2012. Improved moraine age interpretations through explicit matching of geomorphic process models to cosmogenic nuclide measurements from single landforms. *Quaternary Research*, *77*(2), pp.293-304.

Balco, G., Stone, J.O., Lifton, N.A. and Dunai, T.J., 2008. A complete and easily accessible means of calculating surface exposure ages or erosion rates from ^{10}Be and ^{26}Al measurements. *Quaternary geochronology*, *3*(3), pp.174-195.

Balco, G., 2020a. A prototype transparent-middle-layer data management and analysis infrastructure for cosmogenic-nuclide exposure dating. *Geochronology*, *2*(2), pp.169-175.

Balco, G., 2020b. Glacier change and paleoclimate applications of cosmogenic-nuclide exposure dating. *Annual Review of Earth and Planetary Sciences*, *48*, pp. 21-48.

Gosse, J.C., Evenson, E.B., Klein, J., Lawn, B. and Middleton, R., 1995. Precise cosmogenic ^{10}Be measurements in western North America: Support for a global Younger Dryas cooling event. *Geology*, *23*(10), pp. 877-880.

Putkonen, J. and Swanson, T., 2003. Accuracy of cosmogenic ages for moraines. *Quaternary Research*, *59*(2), pp. 255-261.

Young, N.E., Briner, J.P., Schaefer, J., Zimmerman, S. and Finkel, R.C., 2019. Early Younger Dryas glacier culmination in southern Alaska: Implications for North Atlantic climate change during the last deglaciation. *Geology*, *47*(6), pp. 550-554.